Prompt photons at the LHC : selection, measurements of single- and di-photon production cross sections, and Higgs boson searches with the ATLAS detector

Giovanni Marchiori

Université Pierre et Marie Curie

Laboratoire de Physique Nucléaire et des Hautes Energies

Université Pierre et Marie Curie

Habilitation à diriger des recherches

Spécialité : Physique des particules

présentée par

Giovanni Marchiori

Prompt photons at the LHC : selection performance, measurements of single- and di-photon production cross sections, and Higgs boson searches with the ATLAS detector

---

Soutenue le 6 Novembre 2013 devant le jury composé de :

M. Reynald Pain	LPNHE, Paris	Président du Jury
M. Jonathan Butterworth	University College of London	Rapporteur
M. Michel Fontannaz	LPT, Orsay	Rapporteur
M. Vivek Sharma	University of California, San Diego	Rapporteur
M. Louis Fayard	LAL, Orsay	
M. Guillaume Unal	CERN	

à Lucia

Acknowledgements

I would like to thank my wife Lucia for her love and patience and all the support during the past years, even when we lived in different cities or when I was away at SLAC, at CERN or elsewhere, or too busy finalizing some publication – Lucia, this document is dedicated to you, though it can't replace the time I stole from you.

I would also like to thank my family and my friends, for always caring and encouraging and teaching me to do my best, to be curious, and to seek for answers to questions that were never trivial. In particular, this document is also ideally dedicated to the loving memory of my grandfather Battista, one of my greatest supporters – mi spiace che ti sei mancato le belle notizie degli ultimi anni, caro nonno.

I am indebted to all my colleagues and friends in BaBar and in ATLAS, too, especially those in INFN Pisa and LPNHE and those coordinating or participating to the activities of the Breco, SM and SMDP working groups – I learned from all of you, your feedback was very important for my personal growth in HEP.

Finally I wish to express my gratitude to the members of the jury and the three rapporteurs of this habilitation thesis, for their kind availability, for reading and commenting this manuscript and for sharing with me this important moment of my career – I really appreciate that !

Contents

Foreword 9

1 Introduction 11
- 1.1 Physics measurements with prompt photons in the final state of pp collisions . . . 11
- 1.2 The Large Hadron Collider . . . 12
- 1.3 The ATLAS detector . . . 16
  - 1.3.1 The inner tracking detector . . . 16
  - 1.3.2 The calorimeters . . . 18
  - 1.3.3 The muon spectrometer . . . 19
  - 1.3.4 The ATLAS trigger . . . 19
- 1.4 ATLAS data taking . . . 20
- 1.5 Particle reconstruction in ATLAS . . . 20
- 1.6 Photon reconstruction and identification in ATLAS . . . 22
  - 1.6.1 Photon and electron reconstruction . . . 22
  - 1.6.2 Photon direction measurement . . . 25
  - 1.6.3 Photon energy measurement . . . 27
  - 1.6.4 Photon identification . . . 28
  - 1.6.5 Photon isolation . . . 35
  - 1.6.6 Photon triggers . . . 37

2 Photon identification and trigger: performance optimisation and data-driven efficiency estimations 39
- 2.1 Optimization of the photon identification criteria for the first 7 TeV data . . . 39
  - 2.1.1 Previous studies . . . 39
  - 2.1.2 Requirements for the photon identification criteria in the first 7 TeV data . . . 40
  - 2.1.3 Simulated samples . . . 41
  - 2.1.4 Definitions . . . 46
  - 2.1.5 Optimization of the photon identification criteria . . . 47
  - 2.1.6 Results . . . 48
- 2.2 Measurement of the photon identification efficiency with 7 and 8 TeV data . . . 50
  - 2.2.1 Previous studies . . . 51
  - 2.2.2 Measurement of the photon identification efficiency in 2011 data using Z radiative decays . . . 52
  - 2.2.3 Measurement of the photon identification efficiency in 2011 data with the matrix method . . . 55
  - 2.2.4 Final results (2011 data) . . . 57
  - 2.2.5 Photon identification efficiency with 2012 data - preliminary studies . . . 59
- 2.3 Optimization of the photon triggers for the 8 TeV run . . . 60
- 2.4 Measurement of the photon trigger efficiency with 8 TeV data . . . 65

3 Measurement of the production cross sections of isolated prompt photons 71
- 3.1 Theoretical overview . . . 72
- 3.2 Theoretical calculations . . . 77
  - 3.2.1 Inclusive photon and photon-jet production cross section . . . 77

3.2.2 Di-photon production cross section . . . 78
3.3 Measurement of the inclusive production cross section of isolated prompt photons at 7 TeV . . . 79
3.3.1 Event selection . . . 80
3.3.2 Background subtraction . . . 80
3.3.3 Cross section measurement . . . 83
3.3.4 Results . . . 84
3.4 Impact of the isolated photon data on the gluon distribution in the proton . . . 86
3.5 Measurement of the production cross section of isolated prompt photons in association with jets at 7 TeV . . . 88
3.6 Sensitivity of the LHC isolated γ+jet data to the proton PDFs . . . 93
3.7 Measurement of the production cross section of isolated photon pairs at 7 TeV . . . 94

4 Search and discovery of the Higgs boson in decays to final states with photons 101
4.1 Production cross sections and branching ratios for a Standard Model Higgs boson . 102
4.2 Statistical methods . . . 103
4.3 Observation of a Higgs-like boson decaying to di-photons, $H \to \gamma\gamma$ . . . 107
4.3.1 MC-based feasibility studies at $\sqrt{s} = 10$ TeV and 7 TeV . . . 108
4.3.2 Early data analysis: measurement of the backgrounds to the $H \to \gamma\gamma$ search, reappraisal of its sensitivity and first upper limits on the production cross section. . . . 111
4.3.3 Higgs boson discovery . . . 112
4.4 Search for the SM Higgs boson decaying to a photon and a Z boson, $H \to Z\gamma$ . . . 119

Conclusion 127

Bibliography 129

Foreword

My personal adventure in experimental particle physics began in the autumn of the year 2000, when, as an undergraduate student at University of Pisa, I joined the local group of the BABAR collaboration, to prepare my "tesi di laurea" (master degree thesis) on a search for CP violating effects in charmless decays of charged B mesons, under the supervision of Prof. Marcello Giorgi. The B-factory experiments, BABAR and Belle, had just started to take data; for me, that was the beginning of a 10-year engagement in the field of flavor and CP violation physics; for the B-factories, it was the initial period of a successful decade of data-taking that would produce a large number of important results, including the confirmation of the explanation of the origin of CP violation proposed in 1972 by Kobayashi and Maskawa, who extended to three quark families Cabibbo's theory of mixing and were eventually awarded the 2008 Nobel Prize in physics. I still remember with pleasure that period, the work within a group of young and talented physicists, my first trip to SLAC and the shifts in the BABAR control room, the time spent to learn the physics of B mesons and of e^+e^- colliders and to analyse real data to produce my first plots with PAW, ROOT and RooFit(Tools)...

After graduating in July 2001, I was enrolled in the doctoral program "Galileo Galilei" of University of Pisa in January 2002, and, still within the BABAR Pisa group, I started a long-term activity on the measurement of exclusive hadronic B meson decays to open-charm final states, in particular aiming at a measurement of the least-well measured angle, γ, of the $b\bar{d}$ Unitarity Triangle. Under the supervision of Prof. Giovanni Batignani, with the help of Matteo Rama and precious feedback from the conveners and the colleagues of the "Breco" BABAR analysis working group, during the three years and a half that led to my Ph.D. thesis defense in June 2005, I measured the branching ratio of $B \to D^0 K$ decays and performed the first measurement of CP-violating observables in $B \to D^0_{CP\pm} K$ decays [1, 2], to set constraints on γ. In the following years I finalised with a Ph.D. student, Moritz Karbach, the results of the $B \to D^0_{CP} K$ study using the full BABAR dataset (finding one of the first $> 3\sigma$ evidence of direct CP violation in charged meson decays [3]), coordinated several other analyses measuring CP-violating observables in related charged and neutral $B \to D^{(*)} K^{(*)}$ decays, and contributed to the combination of their results in a single, final measurement of γ with $\approx 16^\circ$ uncertainty [4]. The "shepherding" of the γ-related measurements was part of the activities I was involved in, between December 2007 and November 2010, as one of the two conveners of the Breco working group, initially with Vincent Tisserand and later with Vincent Poireau, both at LAPP, in Annecy. The main physics goals of this BABAR analysis working group, which during my convenership included about 15 Ph.D. students and 15 researchers, were the measurement of γ in CP-violating B decays to test the CKM mechanism, and the study of non-CP-violating B decays, for various purposes including searches of new $c\bar{s}$ states, tests of QCD models, search for wide D^{**} resonances, studies of final-state interactions, and investigation of baryon production in baryonic B decays. The convener responsibilities consisted in monitoring and coordinating these analyses, ensuring that the needed resources were available, giving advice, reviewing the progress of the work, the results, and their presentation in international conferences and in publications to be submitted to peer-reviewed journals. During my convenership, 14 new results have been published in either Phys. Rev. D and Phys. Rev. Lett., and – of no less importance – 8 Ph.D. students working on some of these measurements successfully defended their theses.

Beyond data analysis, my training as a physicist during the years of Ph.D. and Post-Doc in Pisa between 2002 and 2008 was integrated by an R&D activity on silicon microstrip detectors for charged particle tracking, either in the cleanroom of the BABAR group to assemble and test

spare modules for the BABAR vertex detector (SVT), at SLAC as the on-call operation manager and data quality responsible of the SVT during data-taking, at the Elettra synchrotron and the Pisa laboratories to measure the charge collection efficiency decrease in radiation-damaged silicon detectors [5], or developing tools (the slow-control system and the condition database-ROOT interface) for the CERN beam test of the SLIM5 monolithic active pixel sensors [6].

Despite the interest and involvement I had in flavor physics and the other aforementioned activities, the responsibilities I had the honor to be assigned, and the group of smart colleagues I was working with, I started to feel around 2008 the need to face new challenges, in a different experiment and a different group. In December 2008 I thus joined the ATLAS group of LPNHE, in Paris, initially as a senior Post-Doc and later (October 2009) as a CNRS "Chargé de recherche". Here, while continuing the BABAR data-analysis activities I was still involved in, I began contributing to the R&D activities of the ATLAS group on the silicon pixel detectors for the future ATLAS upgrades (setting up the new cleanroom for detector characterization, performing device simulation, doping profile measurements, irradiations, beam test data taking and analysis [7, 8, 9]) and to the analysis of the high-energy pp collisions in the ATLAS detector at the LHC, first based on simulated samples and eventually on the real collisions that the LHC started to deliver in spring 2010. As the LPNHE group had been deeply involved in the design and construction of the liquid-argon electromagnetic calorimeter of the ATLAS detector and in the development of the trigger and reconstruction software of electrons and photons, I chose to get involved in the analysis activities of the group, and of the ATLAS Collaboration, that were focused on final states containing prompt photons, *i.e.* photons not originating from hadron decays – an occupation that kept me rather busy, while giving me a lot of satisfaction, until nowadays!

This document, prepared to obtain the "Habilitation à Diriger des Recherches", is a compendium of the photon-related analysis activities I carried on within ATLAS in the past four years and a half. For the sake of internal coherence, I chose photons as the unifying topic and did not describe the other activities (analysis of BABAR data, detector R&D) that were also part of my daily work after defending my Ph.D. thesis. The activities I will describe can be broadly classified into three categories: optimization and/or *in situ* measurement of photon-related performance [10, 11, 12, 13], measurements of the cross sections of Standard Model processes producing prompt photons [14, 15, 16, 17, 18], and searches (leading to discovery!) of a Standard Model (-like?) Higgs boson decaying to final states containing photons [19, 20, 21, 22, 23, 24, 25, 26, 27, 28]. They will be discussed in sequential order in this document, after the initial chapter describing the general feature of the LHC, of the ATLAS detector, and of photon reconstruction and selection in ATLAS; however, they should not be intended as rigid compartments, since work done in one category was often preparatory for another one, or performed in parallel to a second one, and similar techniques were developed for more than a single analysis, trying to reuse as much as possible the expertise gained with one measurement in later, more elaborate studies. To make the document more self-contained than it would have been otherwise, I indulged sometimes in the inclusion of material – concerning for instance the ATLAS detector, the calculation of QCD cross sections, the basic properties of Higgs boson production and decays or the way we deal with the statistical analysis of the data – that may be "standard" and well known to the reader; I hope that in such cases he/she will forgive me for those digressions and skip to the following section.

Chapter 1

Introduction

The main aim of the ATLAS experiment (A Toroidal LHC ApparatuS) [29] at the Large Hadron Collider (LHC) [30] is the investigation of physics above the electroweak symmetry-breaking (EWSB) scale, in an energy range previously unexplored. Expected to be located approximately within 100 and 1000 GeV, this scale becomes experimentally accessible in the highly energetic collisions of multi-TeV proton beams accelerated inside the LHC.

In the Standard Model (SM) [31, 32, 33, 34, 35] of elementary particles, spontaneous EWSB arises from the presence of a Higgs field with a non trivial vacuum structure, characterized by a set of degenerate ground states that minimize the Higgs-field potential and have non-zero expectation value, $v = 246$ GeV. As a consequence of EWSB, the W and Z gauge bosons and the fundamental matter particles (quarks and leptons) acquire a mass, which otherwise needs to vanish for the SM lagrangian without the Higgs field to be renormalizable and gauge-invariant. The quantum excitations of the Higgs field near the physical vacuum correspond to a new fundamental, scalar, electrically neutral particle, the Higgs boson, whose couplings to the other particles are predicted to be proportional to the particle masses [36, 37, 38, 39, 40, 41]. The mass of the new boson must be below 1 TeV in order to ensure that the longitudinal W boson scattering amplitude, which otherwise would grow with the center-of-mass (CM) energy ($\sqrt{s}$) squared, does not violate unitarity for $\sqrt{s} \gtrsim 1$ TeV.

If the Higgs mechanism is not the responsible for EWSB, other physics beyond the Standard Model (BSM) must break the electroweak symmetry and generate the masses of the fundamental particles. Most of the theoretically motivated extensions of the SM predict detectable effects at the TeV scale explored by the LHC, as for instance the existence of new heavy gauge bosons (W', Z'), supersymmetric particles, or gravitons, with masses of the order of 1 TeV.

The LHC also allows the investigation of the behaviour of SM particles in a qualitatively new energy region, where the momentum transferred during a scattering process is large compared to the masses of the W and Z boson and the electroweak symmetry is restored. The comprehensive study of such "hard-scattering" events is thus crucial to test the validity of the Standard Model of the elementary interactions at such high energies or if extensions are needed, and – in case deviations from the SM expectations are observed – to provide insights about these extensions.

In this chapter I will briefly summarize the relevance of physics measurements with photons at the LHC and the main features and performance of the LHC and of the ATLAS detector, with a particular emphasis on the photon reconstruction and selection.

1.1 Physics measurements with prompt photons in the final state of pp collisions

Some key measurements of the ATLAS physics programme involve precise and efficient reconstruction and identification of prompt photons, *i.e.* photons in the final state not originating from hadron decays. The search for a light Higgs boson H, with mass $m_H \lesssim 140$ GeV, decaying to a di-photon final state, $H \to \gamma\gamma$, is an obvious example. Though the predicted branching ratio for such a decay is around 0.23% for $m_H = 125$ GeV, compared to the much larger branching ratios

for decays to $b\bar{b}$ (58%) or $\tau^+\tau^-$ (6.3%) [1], the larger signal-to-background ratio (the production cross sections of the main SM backgrounds at $\sqrt{s} = 14$ TeV are respectively a few μb for $b\bar{b}$ and about five orders of magnitude smaller for $\gamma\gamma$) and the excellent $\gamma\gamma$ invariant mass resolution that can be obtained with modern electromagnetic calorimeters make the di-photon channel one of the best candidates for a light Higgs boson discovery at the LHC [44, 45]. To this purpose, an accurate reconstruction of photons with typical transverse momenta around $40 - 60$ GeV is required.

In case a Higgs-boson candidate is found at the LHC, it becomes fundamental, in order to determine whether the newly discovered particle is the SM Higgs boson or a different object, to fully characterize its properties, including parity, spin, and decays to as many final states as possible. A decay that can be explored at the LHC to provide additional information on the properties of a light Higgs boson is that to a Z boson and a photon, $H \to Z\gamma$. Though the SM predicts a very small signal-to-background ratio in the $Z\gamma$ final state, non-SM scenarios, in which for instance the Higgs boson is a pseudoscalar, a composite particle or is part of a multiplet, lead to potentially large enhancements that can be either observed or ruled out [46, 47, 48, 49, 50]. Compared to $H \to \gamma\gamma$, the reconstruction of $H \to Z\gamma$ decays requires high efficiency for lower p_{T} photons, down to ≈ 15 GeV.

Additional important BSM physics signatures that involve photons in the final state include for instance graviton decays to di-photons in models with extra spatial dimensions [51] and decays of pairs of supersymmetric particles characterized by the production of two energetic photons and large missing transverse energy [52, 53, 54]. Searches for events with highly energetic photons and jets in the final state allow probing BSM physics that may include excited quarks [55, 56, 57], as well as more exotic scenarios like production of quantum black holes [58, 59, 60], quirks [61, 62], and Regge excitations of string theory [63, 64]. In all these cases, photon candidates with transverse momenta above 100 GeV (or even in the multi-TeV range) are searched for.

SM production of photon pairs and of photons in association with jets is the main background for all of these searches. While most of the latter rely on data-driven methods to estimate the SM backgrounds, their selection criteria – especially in the initial phase of the experiment – are optimized on simulated events, to achieve the highest sensitivity to the signal, quantified by the ratio $S/\sqrt{B}$ between the expected number of signal events (S) and the square root of the number of background events (B) in the phase-space region where the final measurement is performed. It is thus important to perform measurements of the SM background cross sections, in order to correctly normalize the simulation of such processes in the optimization of the BSM search techniques.

Finally, prompt photon production at hadron colliders provides a handle for testing perturbative QCD (pQCD) predictions [65] in measurements that do not require to reconstruct hadronic jet and are thus usually affected by smaller systematic uncertainties. Accurate tests of the calculation techniques used to predict the experimental cross sections of other processes, including Higgs boson production, are thus possible. Moreover, as the SM prompt photon production in pp collisions at LHC energies is dominated, at leading order (LO) in the strong coupling constant α_s, by the $qg \to q\gamma$ elementary process (also called "QCD Compton" scattering), measuring the prompt photon production cross section can be exploited to constrain the gluon parton density function (PDF) in the proton [66, 67]. This is extremely useful as Higgs boson production at the LHC is dominated by the (t-mediated) gluon fusion process, $gg \to H$ (which contributes around 87% of the total production cross section for $m_H = 125$ GeV and $\sqrt{s} = 8$ TeV), and the gluon PDF uncertainty thus affects directly the theoretical predictions of the expected Higgs boson signal yield.

1.2 The Large Hadron Collider

The LHC [30] is a two-ring superconducting hadron accelerator and collider, whose tunnel is located between 45 m and 170 m underground, beneath the Franco-Swiss border near Geneva, Switzerland. It was built by CERN (European Organization for Nuclear Research) to provide high-energy proton-proton (pp) collisions, as well as collisions between heavy nuclei (Pb-Pb), to four main experiments. Two of them, ATLAS and CMS (Compact Muon Solenoid), are general-purpose detectors designed to operate at the highest possible collision rate (nominal design luminosity

1. The SM Higgs boson branching fractions and production cross sections in this document are taken from Refs. [42, 43]

$L = 10^{34}$ cm^{-2}s^{-1}), while the other two are dedicated experiments focusing either on precision measurements of the flavor and CP sector of the Standard Model and the search for new physics in rare B meson decays (LHCb), or on the study of strongly interacting matter and the quark-gluon plasma produced in nucleus-nucleus collisions (ALICE). [2]

With a length of 26.7 km, a maximum achieved pp center-of-mass energy $\sqrt{s} = 8$ TeV and a design value of $\sqrt{s} = 14$ TeV, the LHC is the world's largest and highest-energy particle accelerator. It is housed in the tunnel previously used for the LEP e^+e^- collider. Proton bunches circulate in opposite directions in two rings, under the Lorentz force generated by the magnetic field (8.3 T maximum) of 1232, 15 m-long, superconducting dipoles using NbTi windings, kept at cryogenic temperatures (1.9 K) by superfluid liquid helium. Protons are injected into the LHC rings at an energy of 450 GeV from the injector chain Linac2 – Proton Synchrotron Booster (PSB) – Proton Synchrotron (PS) – Super Proton Synchrotron (SPS), and then accelerated to the nominal beam energy. The choice of a proton-proton machine reduces the power consumption due to the large radiative losses typical of electron ring colliders and yields a higher luminosity than $p\bar{p}$ colliders, thanks to the larger bunch intensity obtainable with protons compared to anti-protons. However, this requires two separate magnetic fields to drive the counter-rotating beams; since the LEP tunnel is not large enough to hold two separate proton rings, a twin-bore magnet design has been adopted for the LHC, where two sets of coils and beam channels are located in the same mechanical structure and cryostat.

The LHC project's long history [68] started with a workshop held in March 1984 in Lausanne [69], though approval of the project was given by the CERN Council only 10 years later, in December 1994. Initially planned as a two-stage project to be completed in 2008, it was decided in December 1996 to opt for a single-stage project that would allow commissioning of the machine in 2005. Construction of the LHC components and of its detectors started soon after the approval, while installation was made possible only after LEP was definitely shut down (November 2000) and dismantled (2001-2002). After the assembly of the machine and of the detectors was completed, the LHC was switched on in September 2008, only to be shut down nine days later, following an accident provoked by the overpressure of gaseous helium produced by the heating generated by a faulty super-conducting cable connecting two magnets. Repairs required one year of work and it was decided to limit the beam energy to 5 GeV until further consolidation work would be done in the LHC tunnel during a longer shutdown.

Eventually the LHC started to collide protons at the end of 2009, first at $\sqrt{s} = 900$ GeV, and later at $\sqrt{s} = 2.36$ TeV, exceeding the previous world record of 1.96 TeV held by the Tevatron collider since 2001. In March 2010, after a short winter shutdown, collisions of 3.5 TeV proton beams were successfully established, and then delivered to the experiments throughout the rest of 2010 and 2011, with a short interruption during winter 2010-2011. The instantaneous peak luminosity was progressively increased from about 10^{27} cm^{-2}s^{-1} to 3.65×10^{33} cm^{-2}s^{-1}, and a total of about 5.6 fb^{-1} of integrated luminosity at $\sqrt{s} = 7$ TeV was delivered to ATLAS and CMS. In 2012 the beam energy was increased to 4 TeV and the peak luminosity, constantly above 10^{33} cm^{-2}s^{-1}, rose to 7.7×10^{33} cm^{-2}s^{-1}, close to its design value. A total of about 23.3 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV was delivered to both ATLAS and CMS. At the end of February 2013 the LHC has been shutdown. Almost two years of repair and upgrade activities, concerning both the accelerator and the detector, are planned, in preparation for the run at a center-of-mass energy close to the design value $\sqrt{s} = 14$ TeV and peak luminosity $L = 10^{34}$ cm^{-2}s^{-1}. The evolution of the instantaneous and integrated luminosities delivered by the LHC to ATLAS (and similarly to CMS) are shown in Fig. 1.1 and Fig. 1.2, respectively [70, 71].

A summary of a few characteristic parameters of the LHC pp collisions, for the three data-taking periods (2010, 2011 and 2012) and their corresponding design values, is given in Table 1.1. The beamspot, *i.e.* the three-dimensional (ellipsoidal) distribution of the pp collision points (as determined from the reconstructed event vertices), had typical transverse sizes of 22 μm in 2011 and 15 μm in 2012, and typical longitudinal width of 60 mm in 2011 and 50 mm in 2012 [72].

An illustration of the cross sections of various physics processes in pp and $p\bar{p}$ collisions as a function of $\sqrt{s}$ is given in Fig. 1.3 [73]. The production cross sections for "interesting" events (t,

2. Two smaller experiments are also located in the LHC tunnel, near the CMS and ATLAS interaction points, to measure the total pp cross section via the optical theorem (TOTEM) or the very-forward production of neutral particles (LHCf).

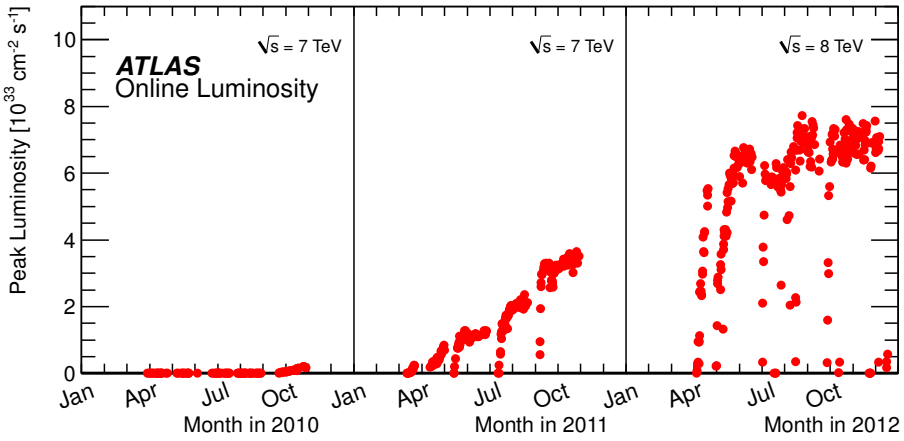

Figure 1.1: The peak instantaneous luminosity delivered to ATLAS per day versus time during the *pp* LHC runs of 2010, 2011 and 2012. The ATLAS online luminosity measurement is used.

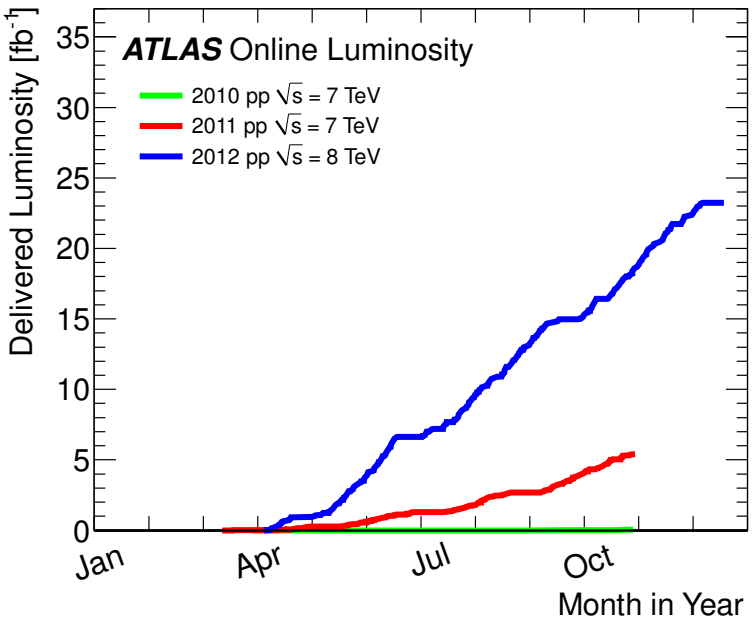

Figure 1.2: Cumulative luminosity versus day delivered to ATLAS during stable beams and for *pp* collisions, for 2010 (green), 2011 (red) and 2012 (blue) running.

Table 1.1: Summary of a few parameters of the LHC proton beam, for three data-taking period and their corresponding design values.

Parameter	2010	2011	2012	nominal
circumference [km]		27		
beam energy [TeV]	3.5	3.5	4.0	7.0
peak instantaneous luminosity (L) [cm^{-2}s^{-1}]	2.1×10^{32}	3.7×10^{33}	7.7×10^{33}	1.0×10^{34}
integrated luminosity ($\int Ldt$) per year [fb^{-1}]	0.048	5.6	23.3	80
number of colliding bunches (n_b) per beam	368	1380	1380	2808
time between collisions [ns]	150	50	50	25
protons per bunch	1.2×10^{11}	1.5×10^{11}	1.6×10^{11}	1.15×10^{11}

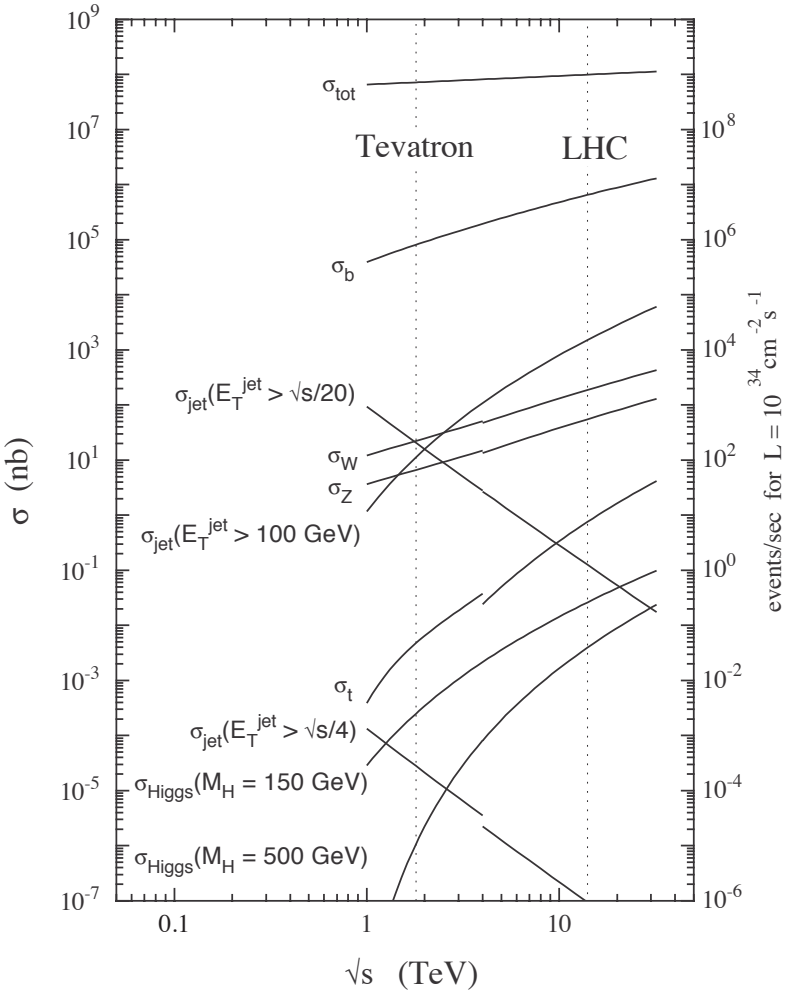

Figure 1.3: Cross sections and rates (for $L = 10^{34}$ cm^{-2}s^{-1}) for various processes in proton-(anti)proton collisions, as a function of the center-of-mass energy.

W/Z, high-p_{T} jets, Higgs boson production, ..) are orders of magnitude smaller than the total cross section, $\sigma_{\mathrm{tot}} \approx 110$ mb at $\sqrt{s} = 14$ TeV. With an inelastic proton-proton cross section $\sigma_{\mathrm{inel}} \approx 80$ mb, at design luminosity $L = 10^{34}$ cm^{-2}s^{-1} the LHC will produce a rate $R = L\sigma_{\mathrm{inel}} \approx 10^9$/s of inelastic events. With a bunch separation of 25 ns, that corresponds to about $\mu = 25$ inelastic events in each bunch-crossing overlapping with the candidate interesting event ("in-time pile-up"). During the 2012 run, at a peak luminosity of 7.7×10^{33} cm^{-2}s^{-1} and a 50 ns bunch separation, the expected number of inelastic cross section per bunch crossing was as high as 40, as shown in Fig. 1.4.

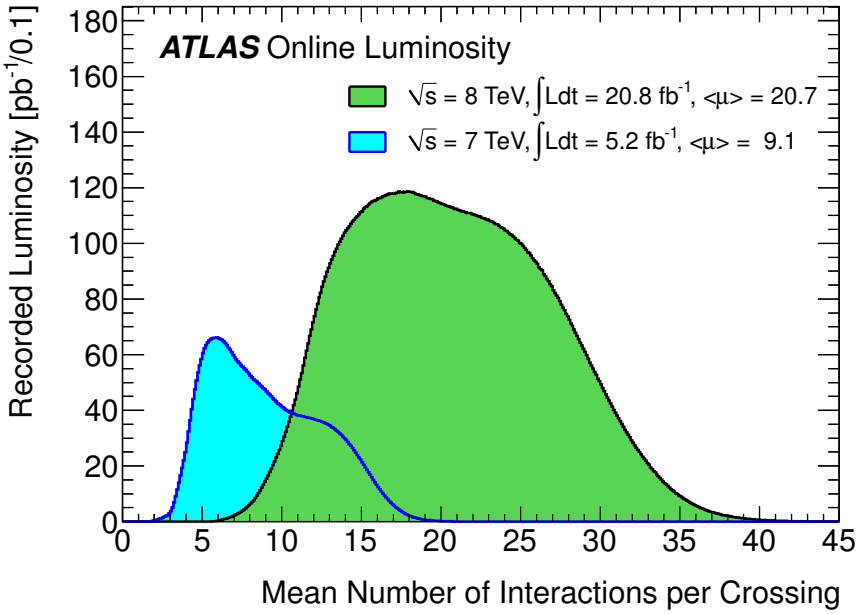

Figure 1.4: Luminosity-weighted distribution of the mean number of interactions per crossing for the full 2011 data and the 2012 data taken between April 4th and November 26th. μ is calculated from the instantaneous per-bunch luminosity $L_b = L/n_b$, the LHC revolution frequency f_r, and the inelastic cross section, $\mu = L_b\sigma_{\mathrm{inel}}/f_r$

1.3 The ATLAS detector

ATLAS [29] is a multi-purpose particle detector with approximately forward-backward symmetric cylindrical geometry [3] and nearly 4π coverage in solid angle. A cut-away sketch of the detector is provided in Fig. 1.5. With a length of 44 m and a diameter of 25 m, ATLAS, which weighs approximately 7000 tons, is the LHC detector occupying the largest volume. It consists of several nested sub-detectors, shown in Fig. 1.6. They are described in the following, starting from the innermost one and going outwards. A particular emphasis is given to the two sub-detectors, the inner tracking detector and the electromagnetic calorimeter, on which photon reconstruction and identification relies.

1.3.1 The inner tracking detector

The inner tracking detector (ID) consists of three subsystems: at small radial distance r from the beam axis ($50.5 < r < 150$ mm), pixel silicon sensors are arranged in three cylindrical layers in the barrel and in three disks in each end-cap; at intermediate radii ($299 < r < 560$ mm), double

3. ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the center of the LHC ring, and the y axis points upward. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$.

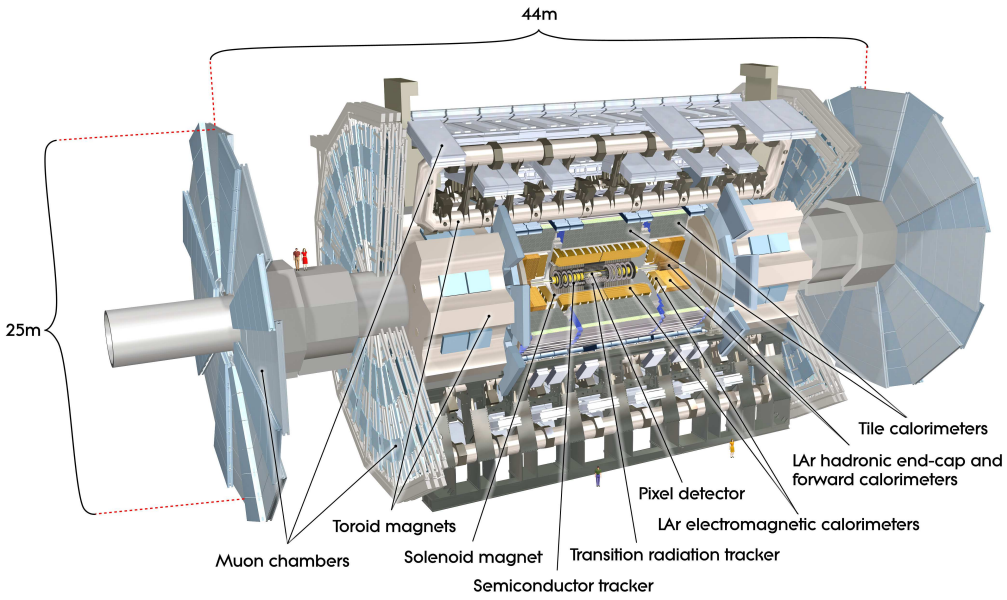

Figure 1.5: Cut-away view of the ATLAS detector.

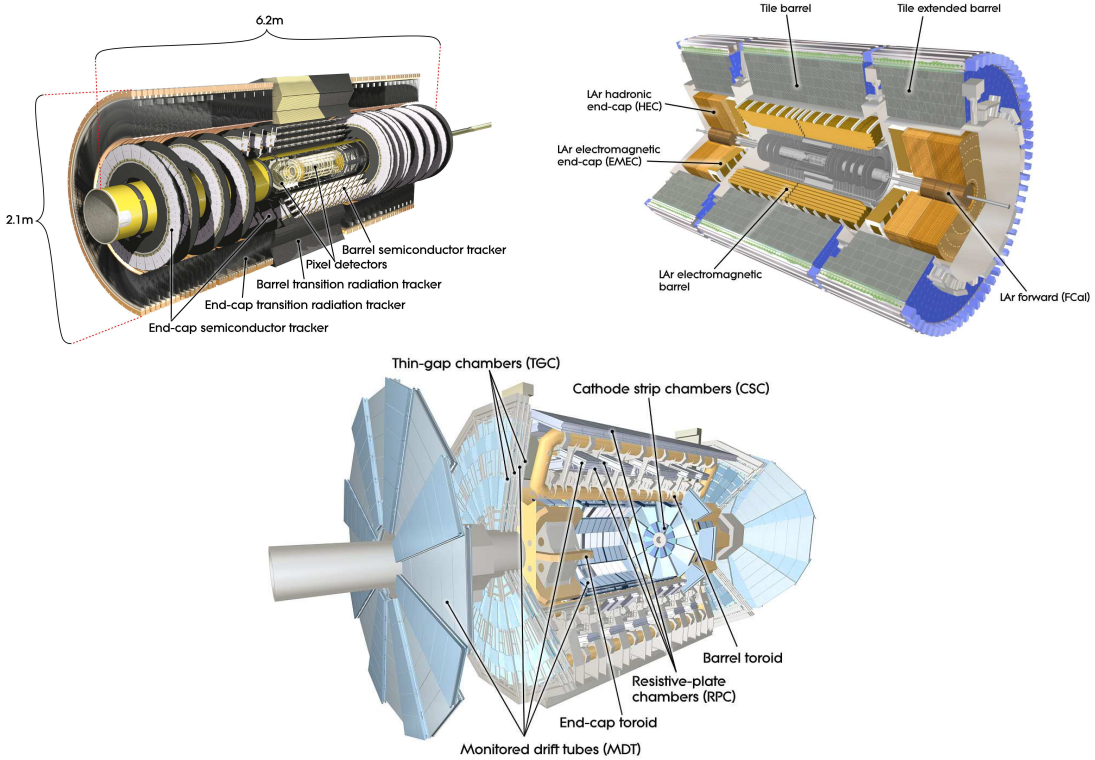

Figure 1.6: Cut-away views of the ATLAS sub-detectors: inner tracking detector (top-left), electromagnetic and hadron calorimeters (top-right), muon spectrometer (bottom).

layers of single-sided silicon microstrip detectors are organized in a system (SCT) of four cylindrical layers in the barrel and nine disks in each end-cap; at larger radii ($563 < r < 1066$ mm), a straw tracker with transition radiation detection capabilities (TRT), divided into one barrel section (with 73 layers of straws, parallel to the beam line and interleaved with fibers) and two end-caps (with 160 layers each of straws radial to the beam line and interleaved with foils), is used. These three systems are surrounded by a thin superconducting solenoid, with a length of 5.3 m and a diameter of 2.5 m, providing a 2T axial magnetic field. The inner detector has full coverage in ϕ. The silicon pixel and SCT subsystems cover the pseudorapidity range $|\eta| < 2.5$, while the TRT acceptance is limited to the range $|\eta| < 2.0$.

The pixel sensors have a minimum cell size of 50×400 μm^2 and intrinsic accuracies of 10 μm and 115 μm along the directions of the short and long sides, respectively. The silicon strips have a pitch of 80 μm, and a small stereo angle (40 mrad) is used in order to measure both coordinates. This yields an intrinsic accuracy of 17 μm and 580 μm on the two coordinates being measured. The TRT only provides $R-\phi$ information in the barrel and $z-\phi$ in the end-caps, with an intrinsic accuracy of 130 μm per straw. However, for the charged particle momentum measurement, the longer trajectory length and the larger number of hits (typically 36 per track) compared to the silicon devices partially compensates the lower intrinsic spatial resolution.

The inner detector allows an accurate reconstruction of tracks from the primary proton-proton collision region, and also identifies tracks from secondary vertices, permitting the efficient identification ("tagging") of b-quark-initiated jets and reconstruction of photon conversions in the inner detector up to a radius of $\approx$ 80 cm. The transition-radiation detection capability also allows electron/pion discrimination.

1.3.2 The calorimeters

A high-granularity lead/liquid-argon (LAr) sampling electromagnetic calorimeter (ECAL) measures the energy and the position of electromagnetic showers with $|\eta| < 3.2$. It is divided into a barrel section, covering the pseudorapidity region $|\eta| < 1.475$, and two end-cap sections, covering the pseudorapidity regions $1.375 < |\eta| < 3.2$. It has three longitudinal sampling layers for $|\eta| < 2.5$ and two for $2.5 < |\eta| < 3.2$. In the region $|\eta| < 1.8$, it is complemented by an additional instrumented argon layer used as a presampler. The ECAL is surrounded by the hadronic calorimeter (HCAL), which measures hadronic showers using either iron-scintillator tiles in the central region ($|\eta| < 1.7$) or LAr sampling calorimeters in the end-cap ($1.5 < |\eta| < 3.2$, copper absorber) and forward ($3.1 < |\eta| < 4.9$, copper-tungsten absorber) regions.

In the following, the electromagnetic calorimeter in the region $|\eta| < 2.5$ is described in more detail; a sketch of a calorimeter module is also shown in Fig. 1.7. The first (also called "front") layer, with a thickness between 3 and 5 radiation lengths (X_0) depending on η, is segmented into high granularity "strips", with width $\Delta\eta$ ranging between 0.003 and 0.006 (with the exception of the regions $1.4 < |\eta| < 1.5$ and $|\eta| > 2.4$), sufficient to provide event-by-event discrimination between single photon showers and two overlapping showers coming from a π^0 decay. The first-layer cell size along the ϕ direction is $\Delta\phi = 0.1$. The second (or "middle") layer of the electromagnetic calorimeter, which collects most of the energy deposited by the photon shower, has a thickness around 17 radiation lengths and a cell granularity of 0.025×0.025 in $\eta \times \phi$. A third ("back") layer, with thickness varying between 4 and 15 radiation lengths and a cell granularity $\Delta\eta \times \Delta\phi = 0.050 \times 0.025$, collects the tails of the electromagnetic showers and provides an additional point to reconstruct the shower barycenter. The presampler is a separate, thin (11 mm) liquid-argon layer, with a read-out granularity 0.025×0.1 in $\eta \times \phi$. The total active thickness of the electromagnetic calorimeter ranges between 22 X_0 and 33 X_0 in the barrel and between 24 X_0 and 38 X_0 in the end-cap, and the longitudinal leakage (*i.e.* the fraction of shower energy that is lost) beyond the end of the active detector of electromagnetic showers initiated by photons and electrons with transverse momentum below 300 GeV is $\ll 1\%$. The absorbers and the electrodes have an accordion geometry. The absorbers are made of lead plates, with thickness ranging between 1.13 mm and 1.7 mm as a function of $|\eta|$. The readout electrodes, located in the gaps between the absorbers, consist of three layers of copper separated by insulating polyimide sheets: the two outer layers are kept at a nominal high voltage and the signal induced on the inner one at ground potential is read out. In the barrel, the high voltage is 2 kV and the typical size of the drift gap on each side of the

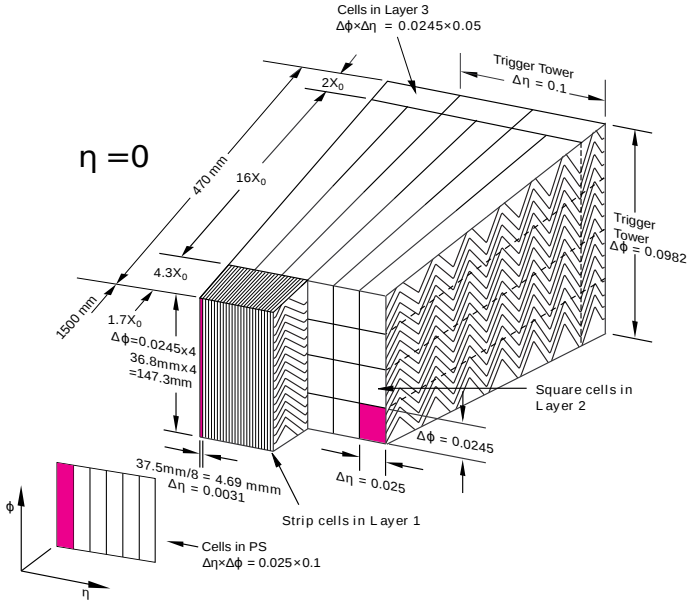

Figure 1.7: Sketch of a barrel module of the ATLAS Pb–LAr electromagnetic calorimeter, where the different layers are clearly visible. The granularities in η and ϕ of the cells of each of the three layers, of the presampler and of the trigger towers are also shown.

electrode is ≈ 2 mm, yielding a total drift time of about 450 ns.

1.3.3 The muon spectrometer

The muon spectrometer (MS) surrounds the calorimeters and consists of three large superconducting air-core toroid magnets, each with eight coils, a system of precision tracking chambers ($|\eta| < 2.7$), and fast tracking chambers for triggering ($|\eta| < 2.4$). The barrel toroid system is 25.3 m long, with inner and outer diameters of 9.4 m and 20.1 m, respectively, and provides a magnetic field of about 0.5 T. Each end-cap toroid is 5.0 m long, with diameters of 1.7 m and 10.7 m, and generates a magnetic field of approximately 1 T. The chambers in the barrel are arranged in three concentric cylinders at radial distances of $\approx$ 5 m, 7.5 m, and 10 m from the beam axis. In the two end-cap regions, muon chambers form large wheels, perpendicular to the z axis and located at distances of $|z| \approx 7.4$ m, 10.8 m, 14 m, and 21.5 m from the interaction point. In the center of the detector ($|\eta| < 0.1$), a gap in chamber coverage allows for services to the inner components of ATLAS. Precision momentum measurement for $|\eta| < 2.7$ is performed by reconstructing the muon curved trajectory in the toroidal field by means of monitored drift tube chambers, except for the innermost end-cap layer, where cathode-strip chambers are used for $2.0 < |\eta| < 2.7$ due to their higher rate capability and time resolution. The spatial resolution of the precision chambers on the coordinate of the muon track in the bending plane is around $35 - 40$ μm. For triggering purposes, resistive plate chambers ($|\eta| < 1.05$) or thin-gap chambers ($1.05 < |\eta| < 2.4$) with time resolution better than 4 ns are used.

1.3.4 The ATLAS trigger

A three-level trigger system selects events to be recorded for offline analysis. The first level ("Level 1", or L1) is based on hardware processors, while the second ("Level 2", or L2) and third ("Event Filter", or EF) levels, collectively called "high-level trigger" (HLT), are implemented through sofware algorithms running on a PC farm. The L1 selects events containing high momentum objects (photons, leptons or jets) from the $\mathcal{O}(10^8)$ proton collisions per second. It operates within a latency of 2.5 μs and produces an output rate of about $50 - 100$ kHz. Event selection is performed using reduced granularity data from the calorimeters and dedicated muon chamber

information. Events passing the L1 trigger are processed by L2, which analyzes the Regions of Interest (RoI) indicated by the L1 and refines the selection, using simplified algorithms that exploit the full detector granularity. The RoIs correspond to about $1-4\%$ of the data of each detector. The L2 has a mean processing time of 40 ms and an output rate of $2-4$ kHz. Events passing the L2 trigger are finally processed by the EF, which employs the same reconstruction algorithms used by offline analyses with only a few exceptions required by the limitations in processing time and produces, within a latency of 4 s, the final output rate of about 400 Hz which is then written to disk for further analysis.

1.4 ATLAS data taking

The ATLAS detector has been collecting data regularly during stable LHC beam collisions. Its average data-taking efficiency during pp collisions amounted to 93.6% in 2010 and 2011, and 93.1% in 2012, leading to a collected integrated luminosity of 44 pb^{-1}, 5.25 fb^{-1} and 21.7 fb^{-1} respectively. The inefficiency accounts for the turn-on of the high voltage of the pixel, SCT and some of the muon detectors (2.0%) and any inefficiencies due to deadtime or due to individual problems with a given subdetector that prevented the ATLAS data-taking to proceed (4.4%). For data analysis, events in which the ATLAS detector is not fully operational, or show data quality problems, are excluded. The efficiency of the data quality requirements was 89.9% in 2011 and 95.8% in 2012 [74], leading to an integrated luminosity of good quality data of 4.7 fb^{-1} at $\sqrt{s} = 7$ TeV and 20.7 fb^{-1} at $\sqrt{s} = 8$ TeV.

1.5 Particle reconstruction in ATLAS

Charged-particle tracks with $p_{\mathrm{T}} > 0.5$ GeV and $|\eta| < 2.5$ are reconstructed by the ID with two complementary track-finding algorithms that start from space points in the silicon detectors and hits in the TRT. The default algorithm starts by searching track seeds in the three pixel layers and extends them throughout the SCT and eventually the TRT. Quality criteria on the track-hit association and on the minimum number of silicon hits are applied, and a final refit of the track with the full information of all the three detectors is performed. Then, to improve the efficiency for secondary tracks from conversions or decays of long-lived particle, a back-tracking algorithm starts from unused track segments in the TRT and extends them inwards to the SCT and pixel detectors. In case no matching silicon hits are found by the back-tracking algorithm, a "TRT-only" track is formed. The track collections produced by the two algorithms are examined to remove ambiguites and double counting and finally merged into a global track collection. The tracking efficiency depends on the particle type, momentum and pseudorapidity, due to multiple scattering, bremsstrahlung (relevant for electrons) and hadronic interactions in the ID material. For $p_{\mathrm{T}} = 5$ GeV it is larger than 99% for muons while it varies between 75% and 95% for electrons and between 80% and 95% for pions as a function of $|\eta|$ [75]. The efficiency is independent of μ within 1%; however, with the default track quality requirements, the number of fake tracks increases with pile-up. An alternative "robust" selection has been also developed for high pile-up conditions, where the fake track rate is minimized and independent of μ, while the tracking efficiency (also independent of μ) is decreased by 5% [76]. The transverse momentum resolution from the simulation is $\frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} = 0.05\% \; \frac{p_{\mathrm{T}}}{\mathrm{GeV}} \oplus 1\%$; the typical resolution for the track direction is $\sigma_\phi \approx 80 \; \mu\mathrm{rad} \left(1 \oplus \frac{45 \text{ GeV}}{p_{\mathrm{T}}}\right)$, $\sigma_{\cot\theta} \approx 1 \times 10^{-3} \left(1 \oplus \frac{7.5 \text{ GeV}}{p_{\mathrm{T}}}\right)$ [75]. The momentum scale and resolution in data agree with the simulation within $\approx 0.1\%$, as determined by comparing in data and Monte Carlo (MC) simulation the invariant mass distributions of several resonances (K^0_S, Λ, J/ψ) decaying to pairs of charged particles [77, 78, 79, 80]. Good data/MC agreement is also observed for the other track parameters.

Electron candidates in the region $|\eta| < 2.47$ are identified by associating charged-particle tracks with deposits of energy in the electromagnetic calorimeters. To account for bremsstrahlung energy loss, electron tracks having enough associated hits in the silicon detectors are fitted using a Gaussian-Sum Filter [81]. The electron four-momentum is formed using the energy measured by the ECAL and the track azimuth and pseudorapidity measured in the inner detector. More details

on the electron reconstruction and energy measurement are provided in Sections 1.6.1 and 1.6.3. The large signal induced on the TRT anode wires by transition radiation X-ray photons and the lateral and longitudinal development of the electromagnetic shower in the calorimeter allow electron/pion discrimination. The TRT alone provides a pion misidentification lower than $\approx 5\%$ for low-p_{T} (< 25 GeV) electrons for an identification efficiency of 90% [75, 82]. The TRT information is combined with shower shape variables of the electromagnetic calorimeter, leakage in the hadronic calorimeter, track quality, track-cluster distance and E/p into a "cut-based" electron identification algorithm, which applies independent requirements (cuts) on each of the input quantities. Three reference sets of cuts have been defined with increasing background rejection power: loose, medium and tight, with an expected jet rejection [4] of about 500, 5000 and 50000 and efficiency around 99%, 95% and 80% for electrons from $Z \to ee$ and $W \to e\nu$, integrated over the range $20 < p_{\mathrm{T}} < 50$ GeV [83].

Photon candidates with $|\eta| < 2.47$ are identified as deposits of energy in the electromagnetic calorimeter that are either not associated to charged-particle tracks or are matched to tracks consistent with a $\gamma \to e^+e^-$ conversion. A detailed description of photon reconstruction, identification, and calibration, is provided in Section 1.6.

Muon candidates are formed starting from tracks reconstructed either in the ID or in the MS [84]. The MS extends the muon reconstruction coverage to the region $2.5 < |\eta| < 2.7$, which is outside of the ID acceptance. If a track is reconstructed both in the ID and in the MS, the two independent momentum measurements are combined ("combined" muons); otherwise the momentum is measured using the MS information ("stand-alone" muons) or the ID information alone. In the center of the barrel region ($|\eta| < 0.1$), which lacks MS coverage, ID tracks are identified as muons using the profile of the associated energy deposits in the calorimeter ("calorimeter"-tagged muons). For $|\eta| > 0.1$, ID tracks not matched to MS tracks are considered muon candidates ("segment-tagged" muons) in presence of hits in the first station of the muon spectrometer nearby the extrapolated particle trajectory. The inner detector tracks associated to muons inside the ID acceptance are required to have a minimum number of associated hits in each of the ID subdetectors to ensure good track reconstruction. The muon reconstruction efficiency is estimated to be around 99% [85] with a tag-and-probe technique using $Z \to \mu\mu$ events in which a kinematic requirements on the di-muon invariant mass and quality criteria on only one ("tag") of the two muons are applied, thus not biasing the second muon ("probe"). The muon momentum resolution $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}}$ increases from around 3% for $p_{\mathrm{T}} = 20$ GeV [86] to $\approx 10\%$ at $p_{\mathrm{T}} = 1$ TeV [84]. The muon momentum scale in data is extracted from the peaks of the $\mu\mu$ invariant mass distributions for Z, Υ and J/ψ di-muon decays: it agrees with MC within $\approx 0.1\%$ [87, 88].

Jets with $|\eta| < 4.5$ are reconstructed starting from three-dimensional, noise-suppressed topological clusters of calorimeter cells, using the infrared- and collinear-safe anti-k_t algorithm [89]. The topological clusters are formed by grouping together cells that have significant energies compared to the expected noise and are adjacent with each other either in the lateral (*i.e.* within the same calorimeter layer) or longitudinal (*i.e.* across different layers) directions. The jet four-momenta are constructed from a sum over their constituent cells, treating each as an $(E, \vec{p})$ four-vector with zero mass. The jet four-momenta are then recalibrated [90] to correct the jet-energy scale for instrumental effects, such as inactive material and non-compensation, as well as for the additional energy due to pile-up. Quality criteria based on the jet timing information and on the fraction of jet energy in the electromagnetic and hadronic calorimeters, as well as on the scalar sum of the transverse momenta of the ID tracks matched to the jet, are applied to suppress fake jets from calorimeter noise, cosmic rays and beam-related backgrounds [90]. For jets in the ID acceptance ($|\eta| < 2.5$), the fraction (jet vertex fraction, JVF) of the scalar sum of the p_{T} of the tracks, associated with the jet and matched to the selected primary vertex, with respect to the scalar sum of the p_{T} of the tracks associated with the jet is sometimes required to be greater than a certain threshold ($0.25 - 0.75$) to reduce the number of pile-up jets. The jet reconstruction efficiency increases with p_{T}, reaching a plateau close to 100% above 25 GeV [90]. The jet energy resolution, $\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$, has a sampling term a of about 60% in the central pseudorapidity region ($|\eta| < 1.5$) and about 100% in the more forward regions; the constant term c is around $2 - 3\%$ for

4. the jet rejection R_j is defined as the inverse of the average probability of a jet to be reconstructed as an electron candidate and to pass the electron identification criteria

$|\eta| < 2.5$ and up to 10% for $|\eta| > 2.5$, while the noise term b is of the order of a few GeV [91, 92]. The absolute jet energy scale is determined *in situ*, using control samples of balanced Z+jet, γ+jet and multijet events [90], with a total uncertainty $< 4\%$ for $|\eta| < 2.5$ and $< 7\%$ for $|\eta| > 2.5$ for jets with $p_{\mathrm{T}} = 40$ GeV [93].

The missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed from energy deposits in the calorimeters and from muon tracks (since muons deposit only small amounts of energy in the calorimeters) [94]. For each calorimeter cell belonging to three-dimensional, noise-suppressed topological clusters, the transverse energy vector $\vec{E}_{\mathrm{T}} = E \sin\theta \hat{n}$ is computed, where E is the measured energy, θ is the polar angle, and $\hat{n}$ is a unit vector, in the transverse plane, pointing from the beam axis to the cell. $E_{\mathrm{T}}^{\mathrm{miss}}$ is the magnitude of the vector opposite to the sum of the $\vec{E}_{\mathrm{T}}$ vectors measured in the calorimeter and of the transverse momenta $\vec{p}_{\mathrm{T}}$ of the reconstructed muon tracks. The calorimeter cell energies are calibrated according to the particle type (photon, electron, hadrons, ..) assigned to the cluster they belong to. As a consequence of momentum conservation and of the quasi-hermeticity of the ATLAS detector, a large missing transverse energy is an indication of high-p_{T} weakly-interacting particles (like neutrinos) produced in the pp collisions. [5] The resolution on the x and y components of the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector is measured using $Z \to \ell\ell$ ($\ell = e, \mu$) and $W \to \ell\nu$ control samples to be around $0.7\ \mathrm{GeV}^{1/2}\sqrt{\sum E_{\mathrm{T}}}$, where $\sum E_{\mathrm{T}}$ is the total visible transverse energy. Suppressing pile-up, either through a jet-area based method [95, 96] or exploiting additional information from the tracking detector, leads to an improvement on the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution of 30% [94, 97].

1.6 Photon reconstruction and identification in ATLAS

In this section, photon reconstruction and performances in the region within the ID and the electromagnetic calorimeter acceptance, $|\eta| < 2.5$, will be described. More details can be found in Refs. [83, 98, 99, 10]. In the transition region ("crack") between the barrel and the end-caps ($1.37 < |\eta| < 1.52$), the performances are expected to be poorer because of the large amount of material upstream of the first active calorimeter layer; the typical performance numbers quoted in the following text do not apply to this region, unless otherwise stated.

1.6.1 Photon and electron reconstruction

In the central region of the ATLAS detector ($|\eta| < 2.5$), photons are reconstructed, together with electrons, by combining information from the electromagnetic calorimeter and the inner tracking detector. Both photons and electrons produce electromagnetic showers in the ECAL that give rise to clusters of neighbouring cells with significant signals; to discriminate between photons and electrons, the presence of a track reconstructed in the inner detector and pointing towards the calorimeter cluster is checked. While traversing the inner detector, photons convert to e^+e^-, with a probability that depends on the amount X/X_0 of material traversed, $P = 1 - e^{-\frac{7}{9}\frac{X}{X_0}}$; when leaving the ID, P varies approximately between 30 and 60% as a function of the photon pseudorapidity, as shown in Fig. 1.8. The resulting ambiguity, in the reconstruction algorithm, between the electron and converted photon hypotheses is resolved by investigating whether the track originates from a conversion vertex or not. In the following the various steps of the photon and electron reconstruction are described in more detail.

- **Seed clusters**
  Photon and electron reconstruction begins with the creation of a set of "seed clusters" of electromagnetic calorimeter cells by means of a sliding-window algorithm. Local maxima of the total transverse energy, deposited in the cells of the presampler and of the three accordion layers contained within a rectangular window of size $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ (corresponding to 3×5 cells of the second layer of the ECAL), are searched for. When a local maximum with a transverse energy above 2.5 GeV is found, a seed cluster is built from all the electromagnetic calorimeter cells contained in the corresponding window. The η and ϕ positions of the seed cluster are calculated as the energy-weighted barycenter of the cells in its core, defined as

5. as the initial longitudinal momentum of the interacting partons is unknown, only the conservation ot momentum in the transverse plane is exploited.

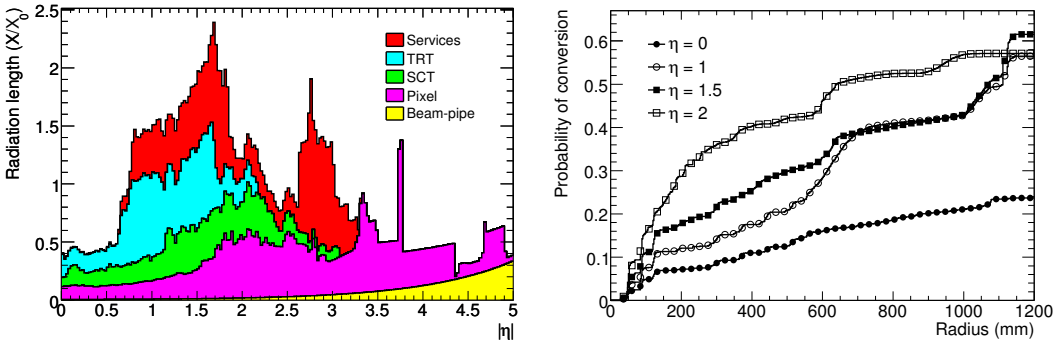

Figure 1.8: Left: material in the inner detector (distance from the beampipe $R < 1150$ mm), in units of radiation length, as a function of $|\eta|$. Right: probability of a photon to have converted as a function of R for different values of pseudorapidity.

the central window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.075$ (corresponding to 3×3 cells of the second layer). If two seed clusters are separated by less than 0.05 along both η and ϕ directions, only the seed cluster with largest transverse energy is kept. The size of the window and the E_{T} threshold are optimized to obtain the best efficiency of the seed cluster search while limiting the rate of fake seed clusters due to noise.

- **Track-cluster matching**
  Seed clusters are initially classified as electrons, converted photons or unconverted photons based on the presence (or absence) of one or more tracks, reconstructed by the inner detector, matching the seed cluster. Track-to-cluster matching is performed by extrapolating the reconstructed track from its last measurement point in the ID and checking whether the η and ϕ coordinates of its expected impact point on the second layer of the ECAL are within a window of size ± 0.05 centered around the cluster barycenter in that layer. To take into account bremsstrahlung energy losses, which are not considered when extrapolating the track to the calorimeter, the size of the matching window is increased to 0.1 in ϕ on the side where the extrapolated track bends as it traverses the inner tracker magnetic field. In case of TRT-only tracks, the matching uses only the ϕ coordinate (and looser criteria) due to the limited spatial resolution provided by the TRT.
  In case of multiple tracks matching the same seed cluster, tracks are ranked according to their reconstruction quality (whether they have hits in the silicon detectors or not) and to the distance $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ between their impact point on the ECAL and the seed cluster position. The track with the better quality and the smallest ΔR is considered as the best match.
  Seed clusters that are not matched to any track are classified as unconverted photon candidates; otherwise, they are considered electron candidates. Seed clusters matched to tracks consistent with originating from a photon conversion in the ID, as described below, are also classified as converted photon candidates.

- **Conversion vertex reconstruction**
  Photon conversions in the ID are reconstructed from pairs of oppositely charged tracks consistent with the hypothesis of being produced in a common origin and with equal directions (collinear with the photon), as implied by the masslessness of the photon. Track pairs are accepted if the difference of their polar angles, the distance between their first hits and their distance of minimum approach are all smaller than a predefined threshold. In addition, the points of minimum approach in both $R - \phi$ and $R - z$ must be close, and the arc length of the $R - \phi$ projection of the two track helices between the line connecting the centers of curvature of the two circles and the actual intersection points must be small.
  For track pairs passing the previous selection, the vertex position and the track momenta at the vertex are estimated from a kinematic minimum-χ^2 fit, based on the measured track

helix-fit parameters (and their covariances) and implementing the common vertex constraint and (via the Lagrange multiplier method) the constraint on the identity of the directions along ϕ and η of the two tracks at the production vertex. The fit is required to converge, with χ^2_{min} and fitted e^+e^- invariant mass not exceeding some predefined threshold. The overall vertex reconstruction efficiency obtained this way, for photons with $p_{\text{T}} = 20$ GeV, is around 80% for photon conversion radii lower than 300 mm.
For radii larger than 400 mm the conversion vertex reconstruction efficiency decreases sharply, as the two tracks are most likely reconstructed by the TRT alone, whose limited spatial resolution does not allow to resolve them and instead a single, "merged" track is reconstructed. A similar effect happens also at smaller radii for highly energetic photons that convert to boosted e^+e^- pairs that are very close in the silicon detectors. For these reasons, and to recover the inefficiency due to "asymmetric" conversions where either the electron or the positron carry a fraction x of the initial momentum of the photon [6] much smaller than one half and falls below the p_{T} threshold (0.5 GeV) required to produce a detectable track in the ID, "single-track" conversions are also reconstructed. These are ID tracks not assigned to any conversion vertex, without hits in the b-layer (the first layer of silicon pixels), and with transition radiation consistent with the electron hypothesis; the conversion vertex position is defined as the position of the first track hit. With the inclusion of single-track conversions, the photon conversion reconstruction effciency is increased to more than 90% over the pseudorapidity region within the TRT acceptance ($|\eta| < 2.0$) for radii extending up to 800 mm.

- **Conversion vertex-cluster matching**
  Single-track conversion vertices are matched to seed clusters using the standard track-cluster matching described previously, extrapolating the associated track from its last measurement point to the second layer of the electromagnetic calorimeter.
  Two-track conversion vertices where the two track momenta differ by less than a factor 4 from each other are considered to be matched to a seed cluster if each track, extrapolated to the calorimeter, matches that same cluster.
  For two-track conversions where the two track momenta differ by more than a factor 4 from each other, a straight-line extrapolation based on the vertex position and the fitted photon direction is used for the cluster matching.
  When multiple conversion vertices are matched to the same cluster, they are ranked: double-track conversion candidates have precedence over single-track ones, and vertex candidates with smaller conversion radius have precedence over candidates with the same number of tracks but larger radius.

- **Photon and electron classification**
  After the initial classification, a calibration (described in Sec. 1.6.3) is applied to compute the cluster energy. To avoid (usually low-p_{T}) fake clusters from (constantly or sporadic) noisy channels, "cleaning" cuts are applied by requiring that the fraction of the reconstructed energies in the presampler and in each layer of the electromagnetic calorimeter do not exceed a threshold close to one. At this stage, a few unconverted photons (order of 10% for photons of a few tens of GeV), whose clusters are erroneusly matched to low p_{T} tracks, and most of the converted photons are (also) reconstructed as electrons. The final arbitration between the electron, converted photon and unconverted photon hypotheses is performed in the following way, based on:

  – the reconstruction quality of the tracks (whether they are TRT-only tracks or have also hits in the silicon detectors),

  – the transverse momentum of the tracks,

  – the compatibility between the track momentum and the energy measured in the cluster,

  – the presence of a hit, associated to the track, in the b-layer.

  Electron candidates are classified as converted photons if they satisfy one of the following

6. the differential cross sectio $d\sigma/dx$ is proportional to $1 - \frac{4}{3}x(1-x)$

conditions:

- the track is TRT-only, $p_{\mathrm{T}} > 2$ GeV and $E/p < 10$ (where the energy E is measured by the calorimeter and the momentum p from the ID);
- the electron cluster is matched to a conversion vertex, and the electron track coincides with a track coming from the vertex (except for two-track conversions when the matched track has a b-layer hit and the other has not);
- the electron cluster is matched to a conversion vertex, the electron track does not coincide with a track from the vertex, and the candidate converted photon p_{T} is larger than the track p_{T}.

Electron candidates are classified as unconverted photons if they satisfy one of the following conditions:

- the track is TRT-only and has $p_{\mathrm{T}} < 2$ GeV
- the electron has not been classified as a converted photon, and the (best) matched track has $p_{\mathrm{T}} < 2$ GeV or $E/p > 10$.

- **Final clusters**
  After the classification of the seed clusters is done, the final electromagnetic clusters are built, with a transverse size that depends on the object classification. In the barrel, a cluster size $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ (corresponding to 3×5 cells in the second layer) is used for unconverted photons, while a size of 0.075×0.175 (3×7 second-layer cells) is used for converted photons, to compensate for the opening between the conversion products in the ϕ direction due to the solenoidal magnetic field, and for electrons, to account for the bending of the electron tracks and the emission of soft bremsstrahlung photons around the electron direction. In the end-cap, where the cell size along θ is smaller than in the barrel and the conversion tracks are closer in ϕ because of the smaller inner radius of the ECAL, a cluster size of 0.125×0.125 (5×5 second-layer cells) is used for all candidates. The choice of the optimum cluster size results from the trade off between the competing requests of limiting the contribution of the electronic noise, of the event pile-up and of other particles in the same physics event (favouring a small cluster size), and of reducing the lateral leakage, *i.e.* the fraction of shower energy lost outside of the cluster (favouring a larger cluster size), since all these quantities, fluctuating on an event-by-event basis, degrade the energy resolution.

The photon reconstruction efficiency as a function of pseudorapidity and transverse momentum, as obtained from a simulated sample of Higgs boson ($m_H = 120$ GeV) decays to photon pairs, is shown in Fig. 1.9. The reconstruction efficiency for photons with $|\eta| < 2.37$, not passing through the calorimeter crack region ($1.37 < |\eta| < 1.52$), and having $p_{\mathrm{T}} > 20$ GeV, is estimated to be close to 98% (greater than 99.5% and 94% for unconverted photons and converted ones, respectively [10]). The inefficiency for converted photons is due to clusters being incorrectly classified as electrons.

1.6.2 Photon direction measurement

For each cluster, the η and ϕ positions are first calculated independently in each layer of the electromagnetic calorimeter, as the energy-weighted barycenters of all cluster cells in the layer. Using simulated samples of single electrons and photons with energies spanning the range $5 - 1000$ GeV, these positions are then corrected for the following systematic biases:

- an η position correction is applied, to account for the bias towards the centers of the cells introduced by the finite size of the readout cells and the small energy sharing between adjacent cells in the η direction. After the correction, the η resolution for 100 GeV photons is $(2.5 - 3.5) \times 10^{-4}$ in the first layer and $(5 - 6) \times 10^{-4}$ in the second layer, fairly independent of the photon pseudorapidity. The same correction for the ϕ position is negligible since the accordion geometry results in more energy sharing between neighbouring cells in the ϕ direction.
- a small bias in the ϕ position, related to the average shower depth with respect to the

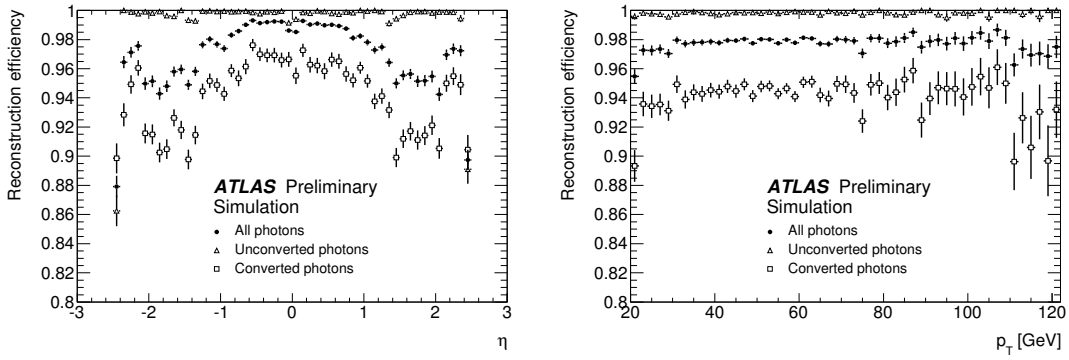

Figure 1.9: Photon reconstruction efficiency as a function of the true photon pseudorapidity (left) and transverse momentum (right), for a simulated sample of Higgs boson decays to photon pairs ($m_H = 120$ GeV, $\sqrt{s} = 7$ TeV). In the right figure, only photons with $|\eta| < 2.37$ and not passing through the calorimeter crack ($1.37 < |\eta| < 1.52$) are considered.

accordion geometry (and thus to $|\eta|$) is corrected. The ϕ position resolution in the second EM calorimeter layer varies within $(5 - 10) \times 10^{-4}$ for 100 GeV photons.

The corrected η and ϕ measurements in the individual layers are then combined to obtain the cluster position. For ϕ, only second-layer information is used (the only combination therefore is done in the overlap region between the barrel and the end-cap). For η, a weighted average of the first and second layer positions is performed, with weights roughly proportional to the inverse of the η resolution in each layer. This implicitly assumes that the incoming particle is projective, *i.e.* its direction is aligned with the nominal origin of the detector reference frame; in that case, the η resolution for 100 GeV photons originated within 5 mm in z from the nominal origin and $|\eta| < 2.5$ (excluding the crack) is $(3 - 4) \times 10^{-4}$. For projective photons, the measured cluster position and the detector frame origin allow the determination of the photon direction. For non-projective photons, however, this estimate is biased. In that case, the photon direction can be estimated using alternative methods:

- using the position of the primary vertex of the hard-scattering event as the photon origin. At high luminosity, the number of reconstructed primary vertices can be large $(20 - 40)$, and robust methods to select the best one are required. While in general hard-scattering events tend to produce high-p_{T} tracks and the primary vertex with the largest scalar sum of the squared transverse momenta of the associated tracks ($\sum_{\mathrm{trk}} p^2_{\mathrm{T,trk}}$) is a sensible choice (and the default one in ATLAS), more refined strategies are required for events with particular topologies with few tracks, like gluon-fusion production of a Higgs boson that decays to a photon pair, $gg \to H \to \gamma\gamma$.
- using the η positions and longitudinal depths of the cluster in the first and second layer of the calorimeter, fitting a straight line in the (R, z) plane through the cluster barycenters in these layers. The resulting resolution on the polar angle of the photon is of the order of $50 - 75$ mrad$/\sqrt{E(\mathrm{GeV})}$. In the search of $H \to \gamma\gamma$, with a calorimeter with a design energy resolution $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E(\mathrm{GeV})}} \oplus 0.7\%$ like the ATLAS one, this angular resolution is below the threshold, 100 mrad$/\sqrt{E(\mathrm{GeV})}$, above which the direction measurement uncertainty starts to give a non-vanishing contribution to the di-photon invariant mass resolution [45].
- for converted photons, using the coordinates of the conversion vertex as an additional point in the straight-line fit mentioned previously.

The second and third methods can also be exploited to extrapolate the photon direction back toward the interaction point and identify, from its intersection with the beam axis, the position of the primary vertex in events, like $gg \to H \to \gamma\gamma$, where the hard-scattering vertex is poorly determined by the recoiling tracks. For instance, using the "calorimeter pointing" technique (second

method), the resolution on the z position of the vertex from each photon is $100-400\ \text{mm}/\sqrt{E(\text{GeV})}$ for photons in the barrel and $700-900\ \text{mm}/\sqrt{E(\text{GeV})}$ for photons in the end-cap.

1.6.3 Photon energy measurement

The energy of both converted and unconverted photons is measured using only the calorimeter cluster information and applying a dedicated energy calibration. The photon energy E is estimated from the energies E_i measured in the cluster cells of each electromagnetic calorimeter layer ($i = 0$ for the presampler, $i = 1..3$ for the three layers of the accordion calorimeter), applying corrections for four effects:

- the energy lost upstream of the calorimeter. The material traversed by particles before reaching the first active layer of the calorimeter (the presampler for $|\eta| < 1.8$ or the first accordion layer elsewhere) has a thickness of $\approx 2-3\ X_0$.
- the sampling fraction, *i.e.* the fraction of energy that is deposited in the active part inside the calorimeter volume.
- the longitudinal leakage.
- the lateral leakage.

The final energy is computed as:

$$E = \Big[\underbrace{a(E_{\text{tot}}^{\text{Acc}},|\eta|) + b(E_{\text{tot}}^{\text{Acc}},|\eta|)\times E_{\text{ps}} + c(E_{\text{tot}}^{\text{Acc}},|\eta|)\times E_{\text{ps}}^2}_{\text{Energy upstream of the accordion calorimeter } (E_{\text{front}})} + \underbrace{\frac{s_{\text{cl}}^{\text{Acc}}(X,|\eta|)}{f_{\text{out}}(X,|\eta|)}\times\left(\sum_{i=1,3}E_i\right)}_{\text{Energy in the accordion calorimeter } (E_{\text{tot}}^{\text{Acc}})} \times \underbrace{(1+f_{\text{leak}}(X,|\eta|))}_{\text{Longitudinal leakage}}\Big] \times \underbrace{F(|\eta|,\phi)}_{\text{Energy modulation}} \tag{1.1}$$

where:

- $a(E_{\text{tot}}^{\text{Acc}},|\eta|)$, $b(E_{\text{tot}}^{\text{Acc}},|\eta|)$ and $c(E_{\text{tot}}^{\text{Acc}},|\eta|)$ are parameters determined as a function of the energy deposited in the three layers of the accordion ($E_{\text{tot}}^{\text{Acc}}$) and of the photon pseudorapidity, $|\eta|$.
- E_{ps} is the part of the cluster energy measured in the presampler, corrected for the fraction deposited in the passive materials.
- X is the the longitudinal barycenter of the shower (shower depth), defined as:

$$X = \frac{\sum_{i=0}^{3} E_i X_i}{\sum_{i=0}^{3} E_i} \tag{1.2}$$

  where X_i is the depth, expressed in radiation lengths, of the longitudinal center of each compartment computed from the center of ATLAS.
- $s_{\text{cl}}^{\text{Acc}}(X,|\eta|)$ is the correction factor to account for the accordion sampling fraction.
- $f_{\text{out}}(X,|\eta|)$ is the lateral leakage correction.
- $f_{\text{leak}}(X,|\eta|)$ is the longitudinal leakage correction.
- $F(|\eta|,\phi)$ is an energy correction that refines the previous corrections (which are based on the $|\eta|$ position of the geometric center of the cell) by taking into account the impact point of a photon inside a cell, which affects the amount of absorber that is traversed (as a function of ϕ) and the lateral leakage.

In the region not instrumented with the presampler ($|\eta| \geq 1.8$) the energy deposited upstream of the calorimeter is parametrized as a function of the shower depth, computed using only the information provided by the three EM calorimeter layers. The calibration coefficients (a, b, c, s_{cl}, f_{out}, f_{leak} and F) are determined from single-photon MC samples covering the energy range between 5 GeV and

1 TeV, $|\eta| < 2.5$ and $|\phi| < \pi$. A similar procedure is used for the calibration of the electron energies. With this procedure, the calorimeter response is expected to be locally uniform within 0.5%. An *in-situ* calibration using $Z \to ee$ decays (and, additionally for photons, $Z \to \ell\ell\gamma$ events) determines the energy scale and intercalibrates the different regions of the calorimeters. $J/\psi \to ee$ decays, which produce lower energy electrons, are used to verify the linearity of the energy response of the calorimeter. The intercalibration procedure determines pseudorapidity-dependent calibration constants $\alpha(\eta)$ (of the order of $\pm 1\%$), that are used to correct the measured photon or electron energy E in data to $E/(1+\alpha)$. Additional cross-checks of the energy scale are performed by inspecting the distribution of the ratio E/p between the energy measured by the calorimeter and the momentum measured by the inner detector for electrons from the more abundant $W \to e\nu$ decays. Both the $Z \to ee$ invariant mass distribution and the $W \to e\nu$ E/p one exhibit peaks that are stable within 0.1% as a function of the number of pile-up interactions [100], thus showing that pile-up has a negligible impact on the photon and electron energy calibration.

The photon energy resolution as a function of the photon energy E (in GeV) is parametrized as:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{1.3}$$

where:

- the "stochastic" or "sampling" term, $a/\sqrt{E}$, is due to eventy-by-event variations in the energy deposited in the liquid argon, induced by fluctuations in the shower development, in the energy lost upstream of the calorimeter and in the lateral and longitudinal leakage of the cluster energy. It is estimated from test-beam data [101, 102], as well as *in-situ* from $J/\psi \to ee$ candidates selected in pp collisions, to be around $10\%/\sqrt{E}$ in the barrel and $15\%/\sqrt{E}$ in the end-cap. The sampling term expected from the simulation agrees within 10% with the value measured in data.
- the "noise" term, b/E, is due to electronic noise of the readout chain of the calorimeter cells spanned by the cluster. Dedicated pedestal runs yield $b \approx 300$ MeV per cluster. The noise term of the resolution is thus only significant for low-energy particles ($E \lesssim 10$ GeV).
- the "constant" term, c, is due to non-uniformities of the calorimeter. Test-beams prior to data-taking and simulations lead to an expected value of $c = 0.7\%$, due to a local constant term below 0.5% over regions of $\Delta\eta \times \Delta\phi = 0.2 \times 0.4$ (ensured by construction tolerances and the electronic calibration system [103]) and an *in situ* intercalibration expected to be better than 0.5%. An "effective" constant term is determined by finding the value of c_{data} such that, after having applied the intercalibration procedure to the data and after adding in the simulation an extra gaussian smearing of the electron energy with size $c_{\text{data}} \times E$ (in addition to a smearing $c_{\text{MC}} \times E$, $c_{\text{MC}} = 0.5\%$), the resolutions of the Gaussian core of the e^+e^- invariant mass distributions of $Z \to ee$ candidates in data and MC samples agree. This procedure yields an effective constant term c_{data} of $1.2\% \pm 0.1\%$ (stat)$^{+0.5}_{-0.6}\%$ (syst) for the barrel and $1.8\% \pm 0.4\%$ (stat)$\pm 0.4\%$ (syst) for the end-caps [83]. This procedure assumes that the MC sampling term reproduces exactly that of the data, and that the non-Gaussian tails of the ee invariant mass distributions in data and MC are also in good agreement; increasing the sampling term by 10% (relative), the effective constant term decreases by about 0.4% (absolute).

1.6.4 Photon identification

The largest background to prompt photon production arises from hadronic jets with a large electromagnetic component, mostly due to the decay to photon pairs of neutral mesons in the jet. In the following they will be also referred to as "fake" photons. The default ATLAS photon identification relies on a cut-based algorithm applying independent requirements to several discriminating variables (DVs) computed from the energy deposited in each of the ECAL cells belonging to the electromagnetic cluster of the photon candidate. The differences between the distributions of such variables for isolated photons and QCD jets reflect the different shapes of their electromagnetic showers: prompt photons typically produce narrower energy deposits in the ECAL and

have smaller leakage in the HCAL compared to fake photons from jets, due to the presence, in the latter case, of additional hadrons near the photon candidate. In addition, fake candidates from isolated $\pi^0 \to \gamma\gamma$ decays – unlike prompt photons – are often characterized by two separate local energy maxima in the finely-segmented strips of the first layer, due to the presence of two photons, as shown in Fig. 1.10.

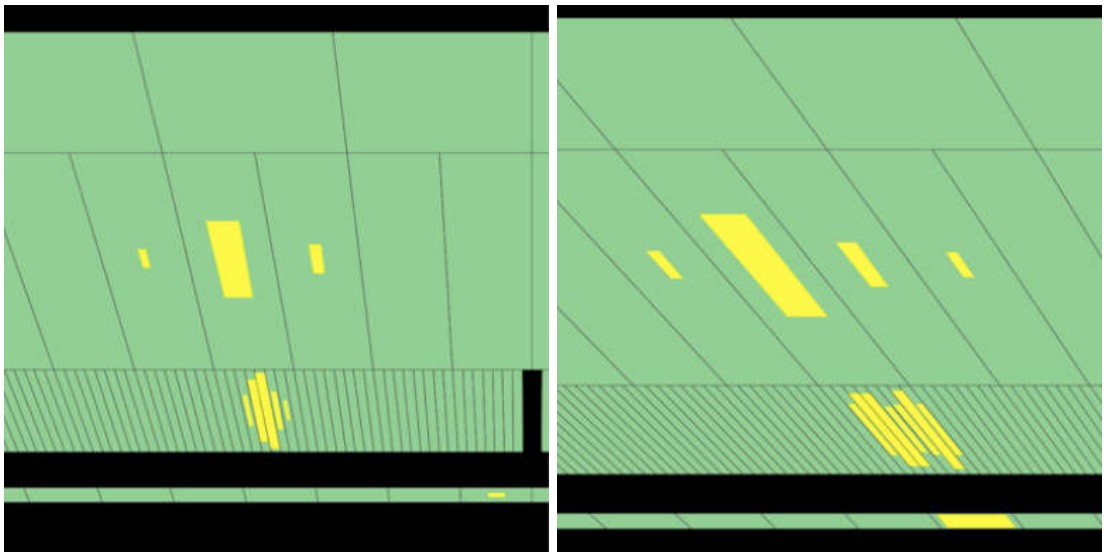

Figure 1.10: Event displays of photon candidates passing (left) or failing (right) tight identification and isolation (see Sec. 1.6.5) requirements, collected at $\sqrt{s} = 7$ TeV. Both displays show a zoom of the calorimeter shower shape around the candidate in the $r - z$ plane. The presampler and the three accordion layers are shown in green; grey lines denote the borders of the readout cells. The yellow rectangles have areas proportional to the transverse energy deposited in the corresponding cell: only cells with a transverse energy greater than 200 MeV are shown. One can clearly see the narrow shower shape in layer 1 for the "photon" and a structure with two peaks from the two close photons from the "π^0" decay.

The definitions of the DVs are the following:

- **Leakage in the hadronic calorimeter**
  Based on the energy deposited in the hadronic calorimeter, the following discriminating variable is defined:
  - *Normalized hadronic leakage*
    $$R_{\text{had}} = \frac{E_{\text{T}}^{\text{had}}}{E_{\text{T}}} \tag{1.4}$$
    is the total transverse energy $E_{\text{T}}^{\text{had}}$ deposited in the hadronic calorimeter, normalized to the total transverse energy E_{T} of the photon candidate.

  In the $|\eta|$ interval between 0.8 and 1.37 the energy deposited in the whole hadronic calorimeter is used, while in the other pseudorapidity intervals only the leakage R_{had_1} in the first layer of the hadronic calorimeter is used.
- **Variables using the second ("middle") layer of the electromagnetic calorimeter**
  The discriminating variables based on the energy deposited in the second layer of the electromagnetic calorimeter are the following:
  - *Middle η energy ratio*
    $$R_{\eta} = \frac{E_{3\times7}^{S2}}{E_{7\times7}^{S2}} \tag{1.5}$$
    is the ratio between the sum $E_{3\times7}^{S2}$ of the energies of the second layer cells of the electro-

magnetic calorimeter contained in a 3×7 rectangle in $\eta \times \phi$ (measured in cell units), and the sum $E^{S2}_{7\times7}$ of the energies in a 7×7 rectangle, both centered around the cluster seed.

– *Middle ϕ energy ratio*

$$R_\phi = \frac{E^{S2}_{3\times3}}{E^{S2}_{3\times7}} \tag{1.6}$$

is defined similarly to R_η. R_ϕ behaves very differently for unconverted and converted photons, since the electrons and positrons generated by the latter bend in different directions in ϕ because of the solenoid magnetic field, producing larger showers in the ϕ direction than the unconverted photons.

– *Middle lateral width*

$$w_{\eta_2} = \sqrt{\frac{\sum E_i \eta_i^2}{\sum E_i} - \left(\frac{\sum E_i \eta_i}{\sum E_i}\right)^2} \tag{1.7}$$

measures the shower lateral width along η in the second layer of the electromagnetic calorimeter, using all cells in a window $\eta \times \phi = 3 \times 5$ measured in cell units.

- **Variables using the first ("front") layer of the electromagnetic calorimeter**
The discriminating variables based on the energy deposited in the first layer of the electromagnetic calorimeter are the following:

– *Front side energy ratio*

$$F_{\text{side}} = \frac{E(\pm3) - E(\pm1)}{E(\pm1)} \tag{1.8}$$

measures the lateral containment of the shower, along the η direction. $E(\pm n)$ is the energy in the $\pm n$ strip cells around the one with the largest energy.

– *Front lateral width (3 strips)*

$$w_{s\,3} = \sqrt{\frac{\sum E_i (i - i_{\max})^2}{\sum E_i}} \tag{1.9}$$

measures the shower width along η in the first layer of the electromagnetic calorimeter, using two strip cells around the maximal energy deposit. The index i is the strip identification number, $i_{\max}$ identifies the strip cells with the greatest energy, E_i is the energy deposit in each strip cell.

– *Front lateral width (total)*
$w_{s\,\text{tot}}$ measures the shower width along η in the first layer of the electromagnetic calorimeter using all cells in a window $\Delta\eta \times \Delta\phi = 0.0625 \times 0.2$, corresponding approximately to 20×2 strip cells in $\eta \times \phi$, and is computed as $w_{s\,3}$.

– *Front second maximum difference.*

$$\Delta E = \left[E^{S1}_{2^{\text{nd}}\max} - E^{S1}_{\min}\right] \tag{1.10}$$

is the difference between the energy of the strip cell with the second greatest energy $E^{S1}_{2^{\text{nd}}\max}$, and the energy in the strip cell with the least energy found between the greatest and the second greatest energy $E^{S1}_{\min}$ ($\Delta E = 0$ when there is no second maximum).

– *Front maxima relative ratio*

$$E_{\text{ratio}} = \frac{E^{S1}_{1^{\text{st}}\,\max} - E^{S1}_{2^{\text{nd}}\,\max}}{E^{S1}_{1^{\text{st}}\,\max} + E^{S1}_{2^{\text{nd}}\,\max}} \tag{1.11}$$

measures the relative difference between the energy of the strip cell with the greatest energy $E^{S1}_{1^{\text{st}}\,\max}$ and the energy in the strip cell with second greatest energy $E^{S1}_{2^{\text{nd}}\,\max}$ (1 when there is no second maximum).

The variables E_{ratio} and ΔE provide rejection against fake photons from $\pi^0 \to \gamma\gamma$ decays when the two showers give separated energy maxima in the first layer, while F_{side} and $w_{s\,3}$ the others provide additional rejection even when the two showers are merged in a wider maximum.

A graphical illustration of the definition of these variables is given in Fig. 1.11.

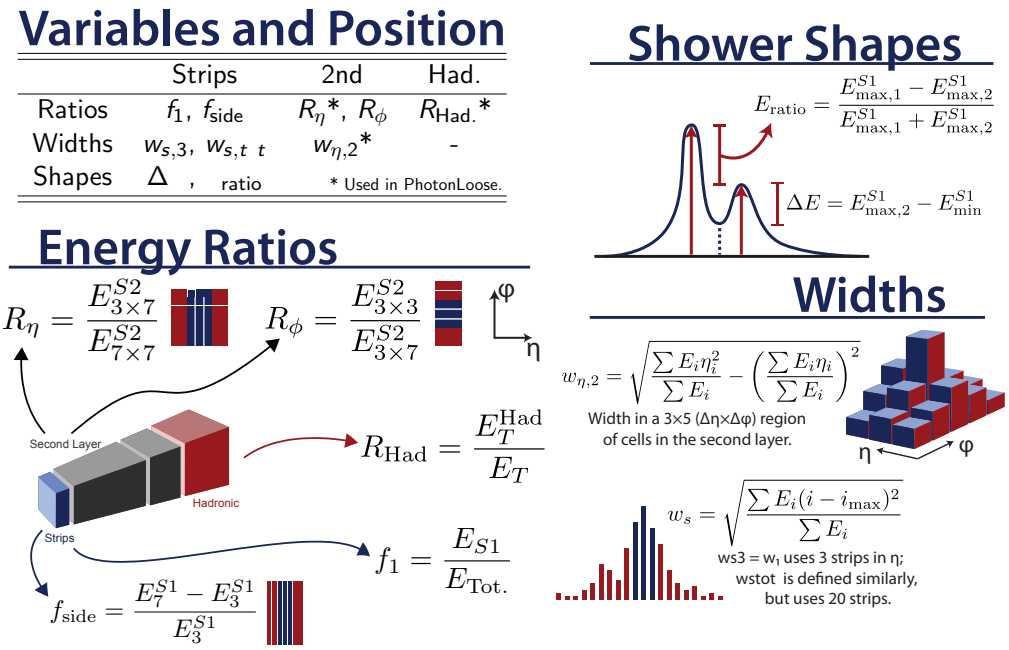

Figure 1.11: Graphical illustration of the calorimeter discriminating variables used for photon identification.

Two reference sets of cuts – "loose" and "tight" – are defined. The loose selection, identical for converted and unconverted photon candidates, is harmonized with the corresponding electron one, and used for triggering purposes. It provides an identification efficiency around 99% for photons with $E_{\text{T}} > 25$ GeV and a hadronic jet rejection factor of about 1000. The tight selection is separately optimized for unconverted and converted photons to provide a photon identification efficiency of about 85% for photon candidates with transverse energy $E_{\text{T}} > 25$ GeV, and a corresponding background rejection factor of about 5000 [10]. The cut-based selection criteria do not depend on the photon candidate transverse energy E_{T}, but vary as a function of the reconstructed pseudorapidity η of photon candidates, to take into account variations of the total thickness of material upstream of the ECAL and of the calorimeter geometry.

Table 1.2 provides a summary of the discriminating variables' definition and whether they are used or not in the loose or tight photon identification. The loose selection includes only shower-shape variables based on information from the ECAL second layer (R_η, w_{η_2}), together with hadronic leakage (R_{had}). As the simulation of the DVs based on the first layer of the calorimeter is rather sensitive to the exact amount of material in front of the calorimeter and to the (large) cross-talk between neighboring cells, these DVs were not considered robust enough to be used for triggering purposes at the beginning of the LHC data-taking and thus the baseline photon triggers do not rely on these variables. Since the distribution of the second-layer variable R_ϕ is rather different between converted and unconverted photons due to the opening in ϕ of the conversion electrons in the ATLAS solenoid field, R_ϕ is not included in the loose selection, in order to ensure a similar trigger efficiency for converted and unconverted photons. The tight selection comprises tighter cuts on the variables used for the loose cut selection, an additional cut on one middle layer quantity (R_ϕ), and cuts on quantities computed from the energy deposit in the strip layer, which – with its fine granularity – provides good $\gamma - \pi^0$ separation. As a consequence, photon candidates

are required to lie in the pseudorapidity region covered by the finely segmented part of the first layer of the electromagnetic calorimeter: photon candidates in the regions $1.37 < |\eta| < 1.52$ and $|\eta| > 2.37$ are thus rejected by the tight identification criteria.

Table 1.2: Variables used for the "loose" and "tight" photon identification algorithms.

Category	Description	Name	Loose	Tight
Acceptance	$\|\eta\| < 2.37$, $1.37 < \|\eta\| < 1.52$ excluded	–	✓	✓
Hadronic leakage	Ratio of E_{T} in the first sampling of the hadronic calorimeter to E_{T} of the EM cluster (used over the range $\|\eta\| < 0.8$ and $\|\eta\| > 1.37$)	R_{had_1}	✓	✓
	Ratio of E_{T} in all the hadronic calorimeter to E_{T} of the EM cluster (used over the range $0.8 < \|\eta\| < 1.37$)	R_{had}	✓	✓
EM Middle layer	Ratio in η of cell energies in 3×7 versus 7×7 cells	R_{η}	✓	✓
	Lateral width of the shower	w_{η_2}	✓	✓
	Ratio in ϕ of cell energies in 3×3 and 3×7 cells	R_{ϕ}		✓
EM Strip layer	Shower width for three strips around strip with maximum energy deposit	$w_{s\,3}$		✓
	Total lateral shower width	$w_{s\,\mathrm{tot}}$		✓
	Energy outside core of three central strips but within seven strips divided by energy within the three central strips	F_{side}		✓
	Difference between the energy associated with the second maximum in the strip layer, and the energy reconstructed in the strip with the minimal value found between the first and second maxima	ΔE		✓
	Ratio of the energy difference associated with the largest and second largest energy deposits over the sum of these energies	E_{ratio}		✓

In Figs. 1.12 and 1.13 the shower shape distributions of photon candidates with $E_{\mathrm{T}} > 20$ GeV in 8 TeV data, as obtained from a high-purity ($\approx 99\%$) sample of radiative Z decays ($Z \to \ell\ell\gamma$, $\ell = e, \mu$), are shown [104]. Details on the selection of this photon control sample, used also for other photon performance studies, are given in Sec. 2.2.2. A calorimeter isolation cut (see Sec. 1.6.5) is also imposed on the photon candidate. The calorimeter shower shape distributions of the photon candidates are compared to those of true photons in simulated $Z \to \ell\ell\gamma$ events and of hadronic jets in simulated $Z(\to \ell\ell)$+jet events, after reweighting their 2D $\{\eta, E_{\mathrm{T}}\}$ distributions to match the data and after correcting the DV values by the average (small) shifts between data and simulation distributions determined from the inclusive sample of isolated photon candidates passing the tight selection per bin of $\{\eta, E_{\mathrm{T}}\}$ and conversion status. A good data/MC agreement for the shower shapes of photons can be observed, as well as the photon/jet discriminating power of the shower shape distributions, even after the isolation requirement, and for the same kinematics.

The description of the *in situ* measurement of the photon identification efficiency and the illustrations of the efficiency as a function of the photon transverse momentum are given in Sec. 2.2.

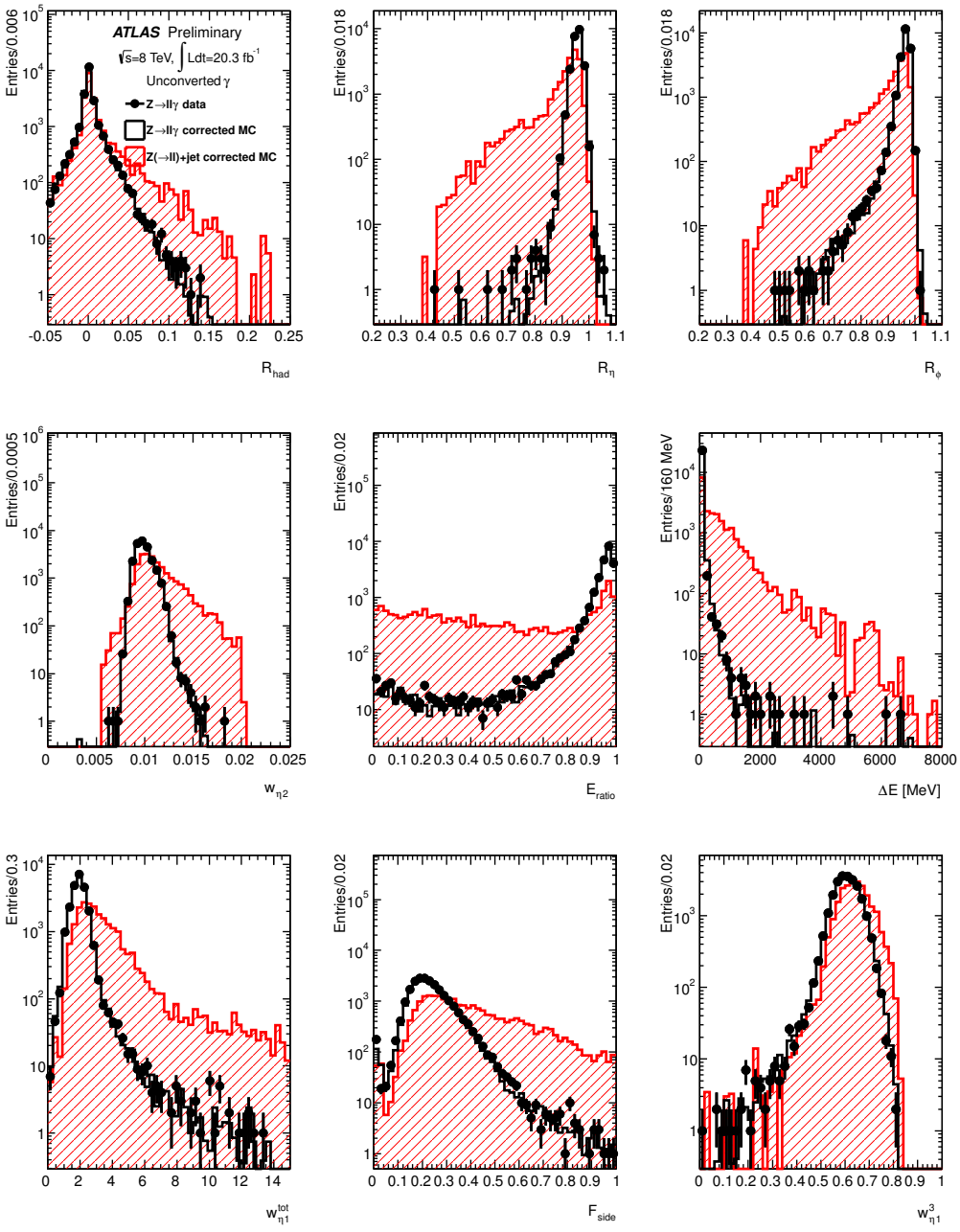

Figure 1.12: Distribution of the calorimetric discriminating variables for unconverted photon candidates with $E_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.37$ (excluding $1.37 < |\eta| < 1.52$) selected from $Z \to \ell\ell\gamma$ events obtained from the $\sqrt{s} = 8$ TeV 2012 data sample (dots). The distributions for true photons from simulated $Z \to \ell\ell\gamma$ events (black hollow histograms) and for fake photons from hadronic jets in $Z(\to \ell\ell)$+jets (red hatched histograms) are also shown, after reweighting their 2D E_{T} vs η distributions to match that of the data candidates. Photon isolation is required on the photon candidate but no criteria on the shower shape are applied. The photon purity of the data sample is $\approx 99\%$.

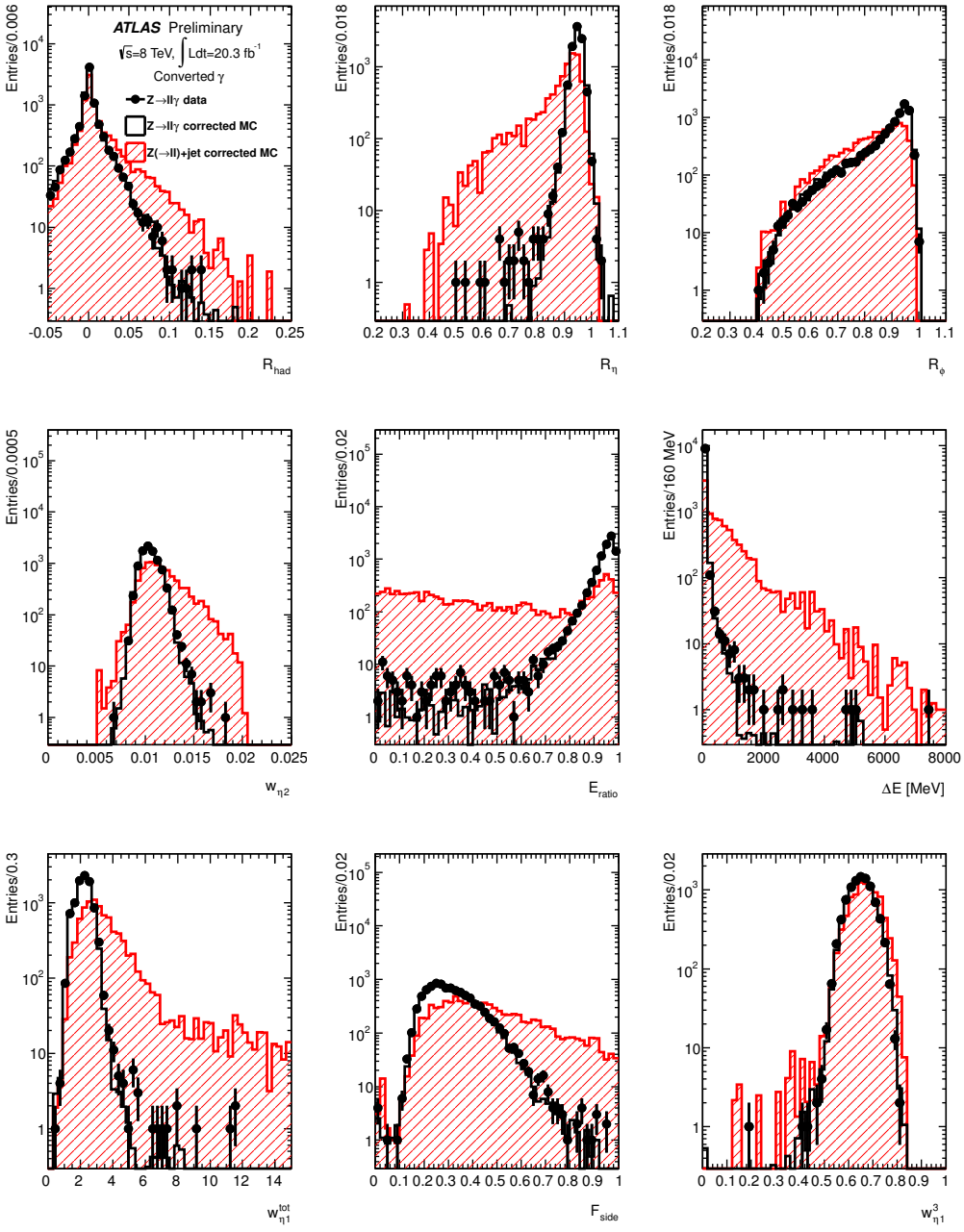

Figure 1.13: Distribution of the calorimetric discriminating variables for converted photon candidates with $E_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.37$ (excluding $1.37 < |\eta| < 1.52$) selected from $Z \to \ell\ell\gamma$ events obtained from the $\sqrt{s} = 8$ TeV 2012 data sample (dots). The distributions for true photons from simulated $Z \to \ell\ell\gamma$ events (black hollow histograms) and for fake photons from hadronic jets in $Z(\to \ell\ell)$+jets (red hatched histograms) are also shown, after reweighting their 2D $E_{\mathrm{T}} vs \eta$ distributions to match that of the data candidates. Photon isolation is required on the photon candidate but no criteria on the shower shape are applied. The photon purity of the data sample is $\approx 99\%$.

1.6.5 Photon isolation

To further suppress the main background from photons originated by the decays of hadrons inside jets, an experimental isolation requirement is usually applied to photons by hadron colliders experiments. The isolation transverse energy ($E_{\mathrm{T}}^{\mathrm{iso}}$), sometimes simply referred to as isolation, is defined here as the sum of the energies deposited in the cells of both electromagnetic and hadronic calorimeters, in a cone (also called "isolation cone") of radius R in the $\eta - \phi$ space around the photon candidate [105]. Larger values of R provide larger photon/jet average separation but lead to isolation variables that are more sensitive to pile-up; the default value for photon-related measurements in ATLAS is $R = 0.4$.

In order to exclude the photon energy from the calculation of $E_{\mathrm{T}}^{\mathrm{iso}}$, the contributions from the electromagnetic calorimeter cells within a window of size 0.125×0.175 (5×7 second-layer cells) in $\eta \times \phi$ centered around the photon barycenter are not included in the sum: the "active" area of the isolation cone is thus $\pi R^2 - 0.125 \times 0.175$. The average value of the small leakage of the photon energy outside this rectangular region, evaluated as a function of the photon transverse energy, is subtracted from the measured value of $E_{\mathrm{T}}^{\mathrm{iso}}$: the typical size of this correction is a few percent of the photon transverse energy. After this correction, $E_{\mathrm{T}}^{\mathrm{iso}}$ for truly isolated photons is nominally independent of the photon transverse energy.

$E_{\mathrm{T}}^{\mathrm{iso}}$ is further corrected by subtracting the estimated contributions from the underlying event and from pile-up. This correction is computed on an event-by-event basis using a method suggested in Refs. [95] and [96]. Based on the standard seeds for jet reconstruction (the noise-suppressed three-dimensional topological clusters), separately for two different pseudorapidity regions ($|\eta| < 1.5$ and $1.5 < |\eta| < 3.0$), a k_T jet-finding algorithm [106, 107] with parameter $R = 0.5$, implemented in `FastJet` [108], is used to reconstruct all jets without any explicit transverse momentum threshold. All positive-energy topological clusters are input to the jet-finding algorithm, and no calibration is performed so that the jet energy remains at the EM scale. Each jet is assigned an area in the following way: every point (η, ϕ) that is within a distance R from any jet is assigned to the closer jet. The transverse energy density for each jet is then computed from the ratio between the jet transverse energy and its area. The ambient transverse energy density due to pile-up and to the underlying event is taken to be the median jet transverse energy density and is multiplied by the active area of the isolation cone to compute the correction to $E_{\mathrm{T}}^{\mathrm{iso}}$. The estimated ambient transverse energy density fluctuates significantly event-by-event, reflecting the fluctuations in the underlying event and pile-up activity in the data. At $\sqrt{s} = 7$ TeV, the mean correction to the calorimeter transverse energy in a cone of radius $R = 0.4$ for an event with one pp interaction is around 500 MeV in simulated events. In the 2010 data, the mean correction is estimated to be 540 MeV for events containing at least one photon candidate with $E_{\mathrm{T}} > 15$ GeV and exactly one reconstructed primary vertex, and increases by an average of 170 MeV with each additional reconstructed primary vertex.

After the leakage and ambient-transverse-energy corrections, the $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution for prompt photons in simulated $qg \to q\gamma$ and $q\bar{q} \to g\gamma$ events is centered near zero, with an RMS width which depends on the radius of the isolation cone and is dominated by the electronic noise in the calorimeter (RMS of 1.5 GeV in 2010 data for $R = 0.4$). An illustration of the isolation distributions for true and fake photons in 2010 data, as determined for the measurement of the SM di-photon production cross section [17], is provided in Fig. 1.14.

The baseline calorimeter isolation variable used for the analysis of 2010 data was computed summing the energies of all the cells inside the isolation cone. In 2011 data, this variable – even after the pile-up correction described before – exhibited a large pile-up dependence, which was understood to originate from the inconsistent treatment of the noise between the raw $E_{\mathrm{T}}^{\mathrm{iso}}$ calculation (a sum of non-noise-suppressed energies of calorimeter cells) and the pile-up correction applied to it (derived from jets built from noise-suppressed topological clusters). Therefore, in the analysis of 2011 and 2012 data, the default photon isolation is computed as the sum of the transverse energies of the subset of cells of the isolation cone belonging to noise-suppressed topological clusters [109]. A graphical illustration of the two isolation definitions is provided in Fig. 1.15.

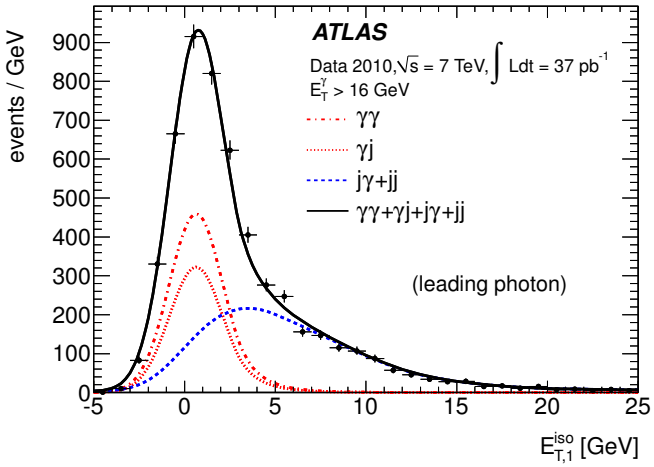

Figure 1.14: Isolation transverse energy of photon candidates with larger E_{T} ("leading" photons) in di-photon candidate events collected in 2010 data [17] (dots with error bars). The red dash-dotted and dotted lines represent the isolation distribution of the true-photon component, either from γj or from $\gamma\gamma$ events. The blue dashed line represents the isolation distribution of the fake-photon component, from $j\gamma$ and jj events.

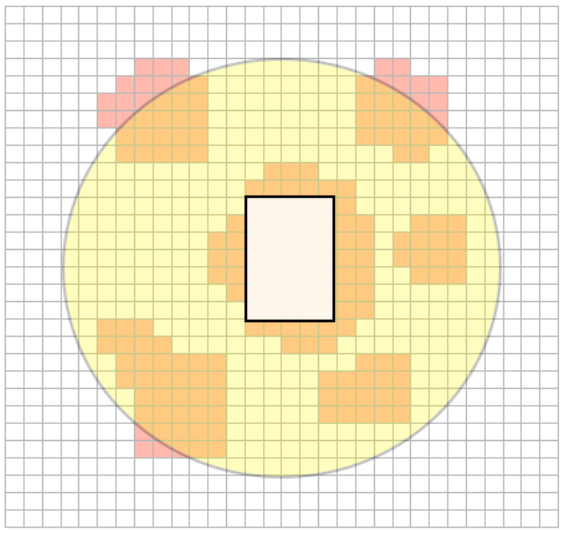

Figure 1.15: Sketch of the calorimeter isolation transverse energy computation. The grid represents the electromagnetic calorimeter middle-cell granularity. The photon candidate energy is mostly contained in the central $\Delta\eta\times\Delta\phi = 5\times7$ rectangle, whose cells are not used for the $E_{\mathrm{T}}^{\mathrm{iso}}$ calculation. A yellow cone of size $R = 0.4$ is drawn around the candidate. In the cell-based $E_{\mathrm{T}}^{\mathrm{iso}}$ variable (no noise-suppression), all cells within this cone are used, whereas in the topological-cluster-based $E_{\mathrm{T}}^{\mathrm{iso}}$ variable (based on topological noise-suppression), only cells in the cone and belonging to topological clusters (orange) are used.

1.6.6 Photon triggers

Events containing photon candidates are selected by the ATLAS three-level trigger system in the following way:

- The L1 trigger uses the signals from the calorimeter cells contained in "trigger towers" to search for possible electromagnetic clusters and compute their transverse energy E_{T} with a precision of 1 GeV. For each trigger tower, which has a granularity of $\approx 0.1 \times 0.1$ in $\eta \times \phi$, the energies of all the cells of the electromagnetic and hadronic calorimeters inside the tower are summed. A sliding-window algorithm is used to identify L1 EM clusters, defined as 4×4 groups of trigger towers that maximize locally the total transverse energy deposited in the cluster towers. A L1 trigger is satisfied if the window's core region of 2×2 trigger towers contains one pair of neighbouring towers with a transverse energy above a programmable threshold, and the 4×4 L1 cluster is the RoI that is further inspected by the high-level trigger.
- At L2, the cells in the second layer of the EM calorimeter within the RoI are inspected to find the one with the largest E_{T}. This "pre-seed" is then used to build a cluster, with algorithms similar to the offline ones. The final cluster position is obtained by calculating the energy-weighted average cell positions on a $\Delta\eta \times \Delta\phi = 0.075 \times 0.175$ grid centered on the pre-seed, while the cluster energy is computed from the energies of the cells within a window $\Delta\eta \times \Delta\phi = 0.075 \times 0.175$ in the barrel and $\Delta\eta \times \Delta\phi = 0.125 \times 0.125$ in the end-cap. Several corrections, used also by the offline reconstruction algorithms, are applied in order to improve the resolution of the cluster position and energy.
- At EF, offline-like algorithms are used for the reconstruction of calorimeter quantities. After retrieving the cell information from a region slightly larger than the RoI, the EF uses the offline sliding-window algorithm to build the cluster and apply all the offline based corrections.

During the 2010, 2011 and 2012 proton-proton collision data-taking periods, the trigger menu has continuously evolved in order to cope with the increasing LHC luminosity. Initially, the trigger relied only on the L1 decision, while the HLT decision was recorded (for performance studies) but not used to reject events. As the luminosity increased, the HLT started to reject events with higher and higher E_{T} thresholds and more stringent selections. For single-photon triggers, based on loose identification requirements, the threshold of the lowest unprescaled trigger was increased from 10 GeV to 30 GeV (while prescaling lower threshold triggers) during 2010, from 60 GeV to 80 GeV during 2011, and it was set at 120 GeV during 2012. For di-photon triggers, based on the same identification requirements applied by single-photon triggers, the threshold was increased from 5 GeV to 15 GeV for both photons in 2010, it was 20 GeV in 2011, and 40 GeV in 2012. Additional di-photon triggers, either with asymmetric E_{T} thresholds (20 and 30 GeV, later increased to 25 and 35 GeV) and loose identification requirements or with symmetric, lower E_{T} thresholds (20 GeV) but stricter identification requirements have also been deployed during the 8 TeV pp run.

Chapter 2

Photon identification and trigger: performance optimisation and data-driven efficiency estimations

Photons produced in physics processes of prime interest at the LHC are expected to have transverse momenta between a few GeV and several TeV. Many of these processes, such as production of a Higgs boson decaying to a photon pair, have small cross sections and suffer from large background, typically from jets of hadrons. It is therefore necessary to have a very efficient photon trigger and excellent photon identification capability, with high photon efficiency and jet rejection, over a broad energy range. The photon trigger criteria have to be optimized in order to find the best compromise between signal efficiency and the maximum acquisition rate set by the DAQ system. The photon identification criteria have to be chosen in order to obtain the maximum jet rejection for a desired photon efficiency, or viceversa.

For both Standard Model measurements and searches for BSM physics involving the reconstruction of prompt photons it is also essential to determine carefully the photon performance of the ATLAS detector. In particular, for SM or Higgs boson cross section measurements in final states with photons, where the experimental results are compared to rather precise theoretical predictions (typically at NLO or NNLO in the strong coupling constant α_s), an accurate knowledge of the photon trigger and identification efficiencies is required in order to limit the experimental systematic uncertainty.

This chapter describes the work I did in the past years on the optimization of the photon identification and trigger requirements and on the *in situ* determination of the identification and trigger efficiencies.

2.1 Optimization of the photon identification criteria for the first 7 TeV data

For the $H \to \gamma\gamma$ search, since the production cross section of the "reducible" di-jet and photon-jet backgrounds are respectively about 6 and 3 order of magnitudes larger than that of the "irreducible" di-photon background, a photon selection with a jet rejection of 5000 or higher is desirable to suppress the reducible background well below the irreducible one [45]. In this section it will be shown how such a jet rejection is achieved, while keeping a large photon efficiency, through requirements on the calorimeter shower-shape variables introduced in Sec. 1.6.4.

2.1.1 Previous studies

In the studies performed on simulated jet and photon samples before the expected LHC start of 2008, the rejection of the photon selection was estimated to be about 5000 for hadronic jets with $p_{\mathrm{T}} > 25$ GeV using only information from the calorimeter DVs and about 8000 exploiting also an

isolation requirement, after optimizing the thresholds for the DV and the isolation selections [98]. To obtain these results, the same set of nine variables listed in Table 1.2 were used, with two exceptions:

- in the range $0.8 < |\eta| < 1.37$, R_{had_1} was used instead of R_{had},
- the rescaled second maximum in the EM strip layer, $R_{\mathrm{max}\,1} = \frac{E^{S1}_{2^{\mathrm{nd}}\,\mathrm{max}}}{1\ \mathrm{GeV}+0.009\,E_{\mathrm{T}}}$, was used instead of the variable E_{ratio}.

The same selection criteria were chosen for unconverted and converted photon candidates. The criteria had been tuned, within the fiducial region $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$, in six pseudorapidity intervals to reflect the pseudorapidity dependence of the DVs as a consequence of both the varying detector readout granularity and the varying amount of material in front of the electromagnetic calorimeter. The criteria were further refined, for $E_{\mathrm{T}} > 20$ GeV, in eight intervals of the photon transverse energy, defined (in GeV) as:

$$[20,25) \quad , \quad [25,30) \quad , \quad [30,40) \quad , \quad [40,50) \quad , \quad [50,60) \quad , \quad [60,70) \quad , \quad [70,80) \quad , \quad [80,+\infty)$$

The study was performed using photons from simulated $H \to \gamma\gamma$ decays ($m_H = 120$ GeV) and jets from a pre-filtered sample containing the simulation of all relevant hard-scattering QCD $2 \to 2$ processes with $\hat{p}_{\mathrm{T}} > 15$ GeV. Both samples were generated with PYTHIA 6 [110], a leading-order parton-shower MC generator that accounts for QED radiation emitted off quarks in the initial state (ISR) and in the final state (FSR), simulates the underlying event using the multiple-parton interaction model, and uses the Lund string model for hadronisation [111]. [1] The filter applied to the jet sample at generator level mimicked the L1 trigger criteria, requiring the summed transverse energy of all stable particles (excluding muons and neutrinos, that deposit small or negligible energy in the calorimeter) in a region of $\Delta\phi \times \Delta\eta = 0.12 \times 0.12$ to be above 17 GeV. The use of such filter permitted to fully simulate only a fraction of events (around 7%) where a jet could potentially be reconstructed as a photon candidate with $E_{\mathrm{T}} > 20$ GeV, thus reducing significantly the computing time needed to produce a large enough data set to study selections with a jet rejection around 5000. The identification efficiency was estimated to be about 85% for photons with $E_{\mathrm{T}} > 25$ GeV and pseudorapidity within the fiducial region.

2.1.2 Requirements for the photon identification criteria in the first 7 TeV data

Following the LHC accident in September 2008, the start of the multi-TeV data-taking was delayed to spring 2010, with a lower center-of-mass energy than initially planned (7 TeV instead of 14). Meanwhile, during 2009, the ATLAS reconstruction, simulation and trigger software continued to be updated; at the same time, the various ATLAS working groups focusing on measurements of final states with photons converged on the following official set of requirements for the photon identification criteria for the first run at 7 TeV:

- to allow efficient identification ($\gtrsim 85\%$ for reconstructed candidates with $E_{\mathrm{T}} > 20$ GeV) for photons from $H \to \gamma\gamma$ but also from Standard Model processes, as SM cross section measurements were one of the main goals of the first year of data-taking, in order to commission the detector.
- to provide a rejection ≈ 5000 for QCD jets with $E_{\mathrm{T}} > 25$ GeV.
- to yield similar efficiencies for unconverted and converted photons, in order to reduce the systematic uncertainty on the cross section measurements originating from the limited knowledge of the detector material upstream of the calorimeter and thus of the photon conversion probability.
- to be tighter than the trigger requirements, in order to avoid biases that might be difficult to estimate with the simulation in the early stages of the data-taking when not enough data

1. Unless otherwise noted, all ATLAS MC samples described in this document are passed through a detailed GEANT4 [112] simulation of the detector geometry and response and then reconstructed with the same algorithms used for data. More details on the ATLAS event generation and simulation infrastructure are provided in Ref. [113].

was available for *in situ* measurements of the trigger efficiency.

In addition, in order to maintain the photon selections as simple as possible and to avoid introducing sharp discontinuities in the photon efficiency vs E_{T} curves, that might create artificial bumps in di-photon or photon-jet invariant-mass spectra to be inspected in the searches for BSM resonances, it was preferred to use the same threshold over the full E_{T} range instead of using E_{T}-dependent discrete (*i.e.* binned) cut thresholds. It became thus necessary to update the previous studies and perform a re-optimization of the thresholds of the DV selection criteria. That work, which I did with Valeria Perez-Reale during 2009 within the e/γ ATLAS working group and whose results are summarized in [10], is presented in the following sections.

2.1.3 Simulated samples

For the optimization and evaluation of the performance of the photon identification requirements, several simulated samples are used:

- a "signal" sample of photon candidates, reconstructed in a sample of 4 million γ-jet (γj) events, matched to true prompt photons. In this sample, prompt photons are (mostly) produced by the leading order Feynman diagrams $qg \to q\gamma$ and $q\bar{q} \to g\gamma$.
- a "background" sample of photon candidates, reconstructed in a jet-filtered sample of 10 million QCD di-jet (jj) events, not matched to true prompt photons.
- additional samples of reconstructed true photons from $H \to \gamma\gamma$ decays ($m_H = 120$ GeV, 100 thousand events) and Randall–Sundrum graviton [51] $G \to \gamma\gamma$ decays ($m_G = 500$ GeV, 300 thousand events) are also used to evaluate the selection efficiency in the medium ($40 - 100$ GeV) and high (> 100 GeV) E_{T} range, respectively.
- the background sample is complemented by a high-statistics unfiltered di-jet sample, generated with the same settings as the filtered one except for the filter requirement and not passed through the ATLAS detector simulation. This sample is used to estimate the average generator-level jet multiplicity in order to normalize the jet rejection factor, as described below.

A minimum transverse energy of 20 GeV is required for both signal and background candidates. All the samples are generated using Pythia 6, at a center-of-mass energy of 10 TeV. The latest versions of the ATLAS event generation and simulation software available by the time of this study are used, to benefit from the latest changes in reconstruction algorithms and improvements in modeling of the detector material, cross-talks effects in the signal digitization, and so on.

Reconstructed photon candidates and true particles are associated in the following way:

- all "final-state particles" (non-decaying particles produced by the generator, excluding neutrinos, with $p_{\mathrm{T}} > 1$ GeV) are extrapolated to the inner surface of the second layer of the electromagnetic calorimeter, and the coordinates of the impact point $(\eta^{\mathrm{extr}}, \phi^{\mathrm{extr}})$ are computed;
- for each reconstructed photon candidate, the distances in η and ϕ, $\Delta\eta = \eta^{\mathrm{extr}} - \eta^{\mathrm{clus}}$ and $\Delta\phi = \phi^{\mathrm{extr}} - \phi^{\mathrm{clus}}$, between each true particle impact point and the photon cluster barycentre in the ECAL second layer, $(\eta^{\mathrm{clus}}, \phi^{\mathrm{clus}})$, are used to find the best-associated true particle:
  - inside an ellyptical cone $(\Delta\eta/0.025)^2 + (\Delta\phi/0.05)^2 < 1$, the true photon with highest p_{T} is chosen; if no photons are found, the true particle with highest p_{T} is selected.
  - if the previous criterion fails, the true particle with the smallest $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ inside a cone $\Delta R < 0.1$ is selected.

Hadron jets faking photons are separated in quark–initiated and gluon–initiated jets, depending on the type of the highest-E_{T} parton from the generator record inside a cone of $\Delta R = 0.4$ around the reconstructed photon candidate.

To illustrate the discriminating power of the shower shape variables, Fig. 2.1 shows the behaviour of the mean of each calorimetric DV as a function of the pseudorapidity $|\eta|$, for both true

(from γj events) and fake (from jj events) reconstructed photon candidates with $E_{\mathrm{T}} > 20$ GeV, before any selection. Figures 2.2 and 2.3 show as an example the normalized distributions of the

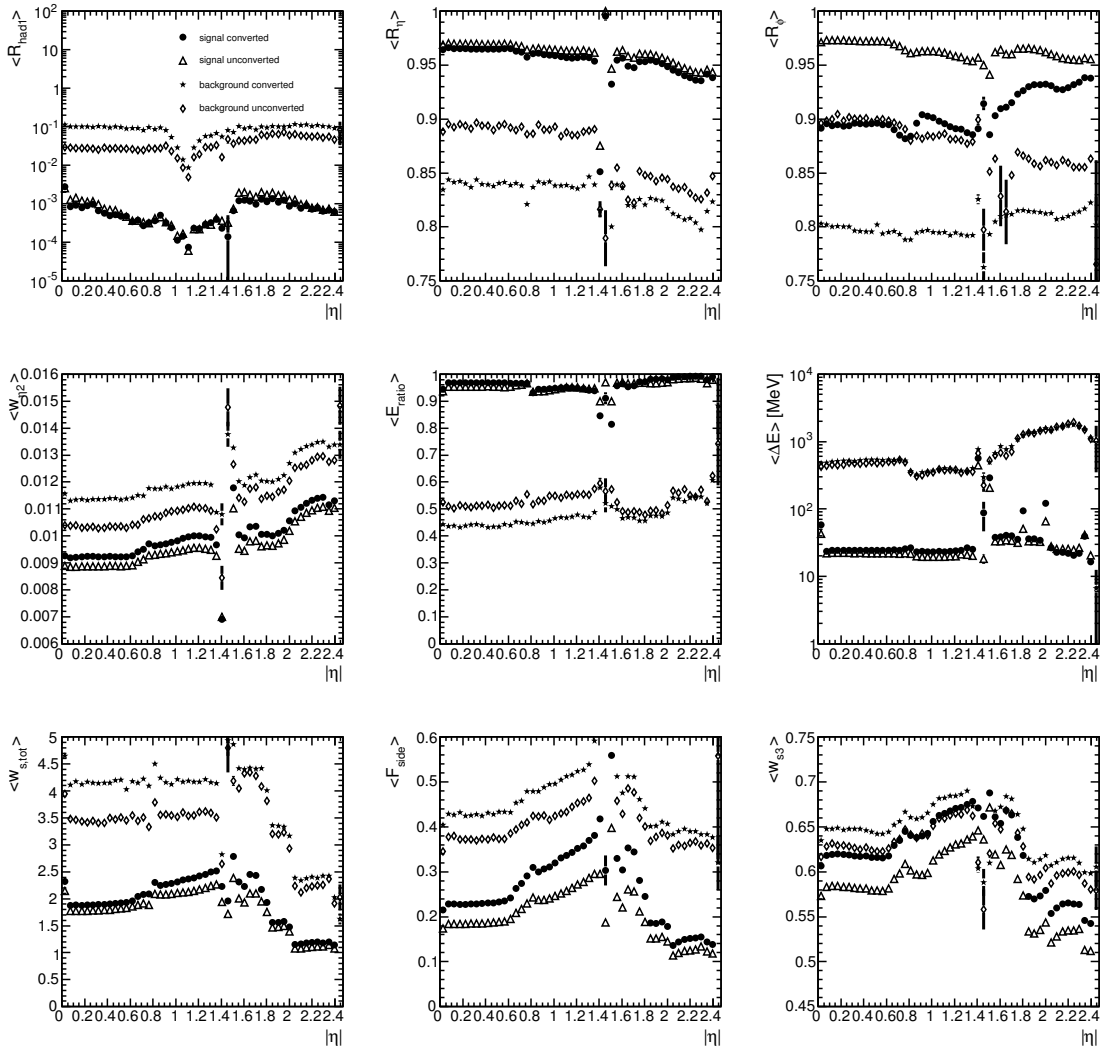

Figure 2.1: Distributions of the means of each calorimetric discriminating variables as a function of the pseudorapidity $|\eta|$ for true and fake photons with $E_{\mathrm{T}} > 20$ GeV before any selection. Both true and fake photons are separated in converted and unconverted candidates. Because of the systematically smaller average value of R_{had_1} in the region $0.8 < |\eta| < 1.37$ (top left plot), the normalized total hadronic energy R_{had} is used instead in this zone to discriminate hadronic showers (see text for details).

calorimetric discriminating variables in the region $0 < |\eta| < 0.6$ for $E_{\mathrm{T}} > 20$ GeV for true and fake photons before any selection for candidates reconstructed as unconverted or converted, respectively.

Being based on the same input quantities (the energy deposited in the cells of the electromagnetic cluster), the discriminating variables have non-negligible correlations, as can be seen in Fig. 2.4, which illustrates the correlation matrices of the calorimetric variables for unconverted or converted, true or fake photon candidates with $|\eta| < 0.6$ and 25 GeV $< E_{\mathrm{T}} <$ 40 GeV.

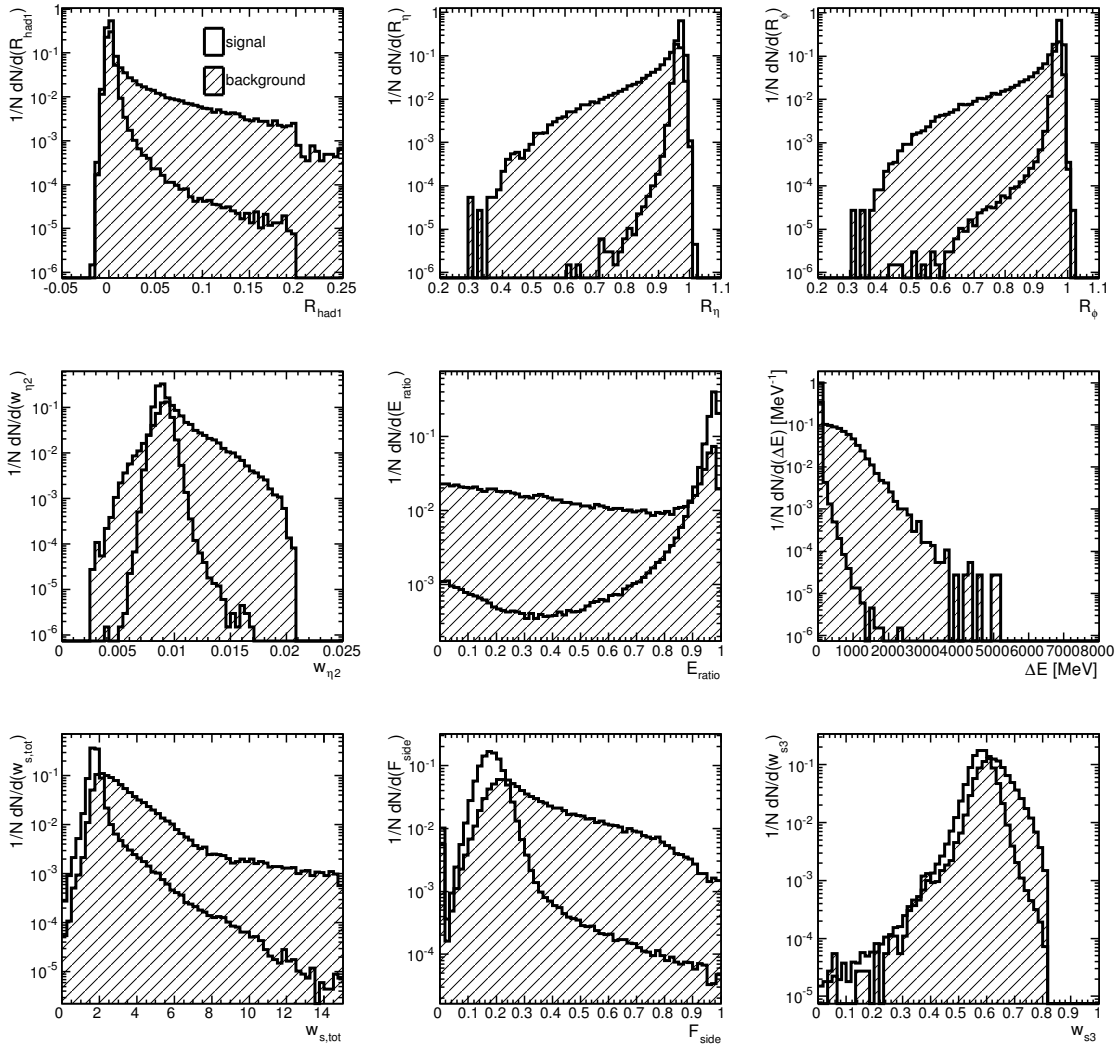

Figure 2.2: Normalized distributions of the calorimetric discriminating variables in the region $0 < |\eta| < 0.6$ for $E_{\mathrm{T}} > 20$ GeV for true and fake photons reconstructed as unconverted before any selection.

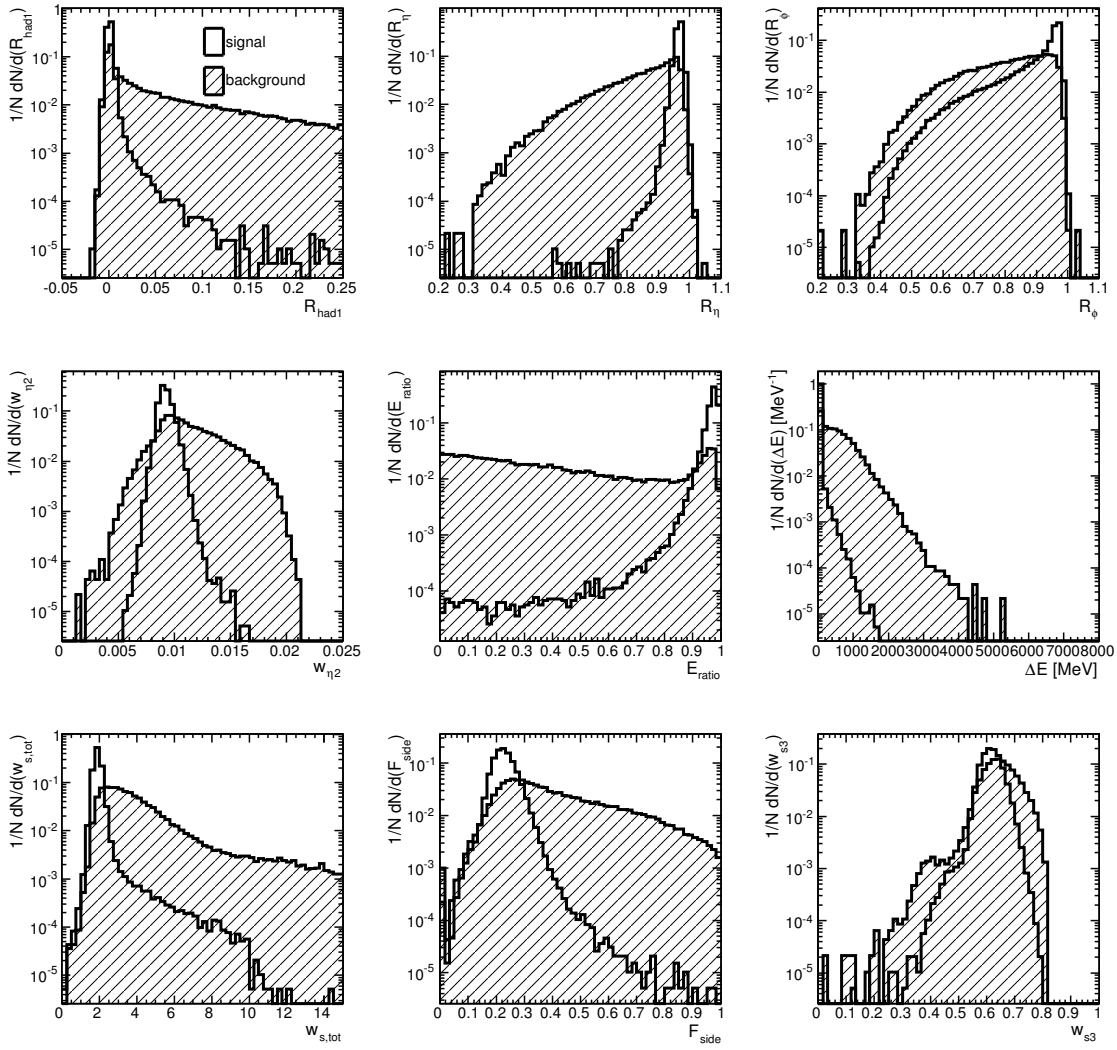

Figure 2.3: Normalized distributions of the calorimetric discriminating variables in the region $0 < |\eta| < 0.6$ for $E_{\mathrm{T}} > 20$ GeV for true and fake photons reconstructed as converted before any selection.

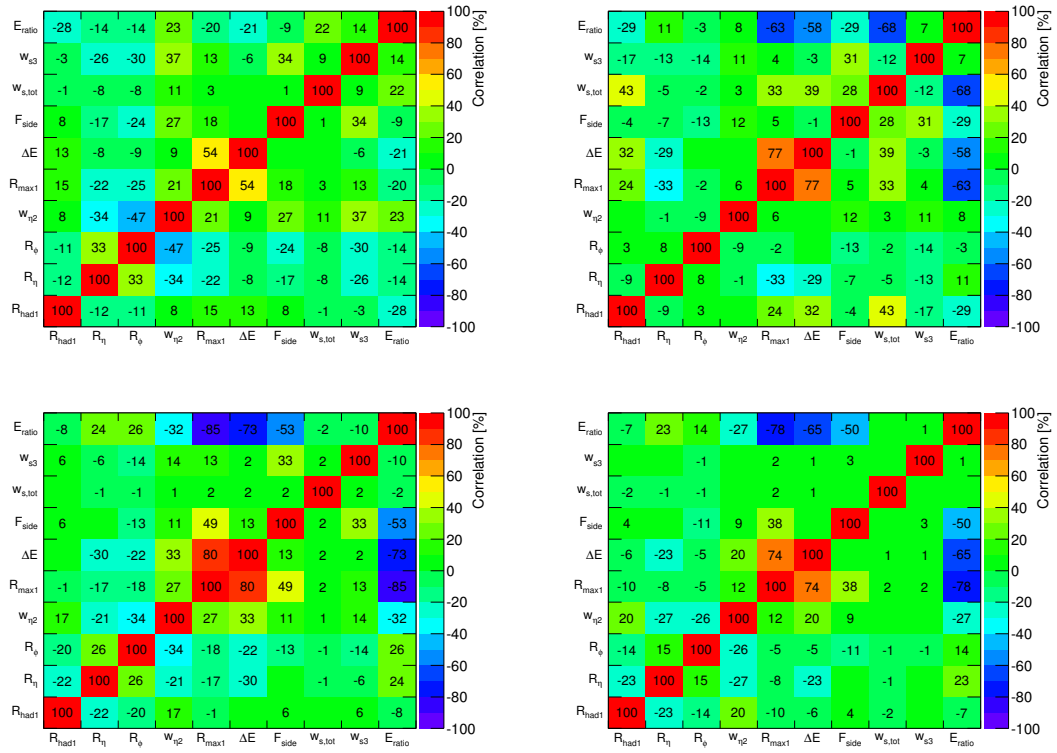

Figure 2.4: Correlation matrices of the discriminating calorimeter variables for unconverted (left column) and converted (right column) photon candidates in the signal (top row) and background (bottom row) samples used for the optimization of the tight selection cuts. The photon candidates have $|\eta| < 0.6$ and 25 GeV $< E_{\mathrm{T}} <$ 40 GeV.

2.1.4 Definitions

Using the available simulated samples, the photon identification efficiency with respect to the generated true photons is computed as:

$$\varepsilon_\gamma = \frac{N_\gamma^{\text{truth, pass cut}}}{N_\gamma^{\text{truth}}}\,, \tag{2.1}$$

where N_γ^{truth} is the total number of true photons having true E_{T} greater than a given threshold (20, 25 or 40 GeV), and $N_\gamma^{\text{truth, pass cut}}$ is the number of reconstructed photon candidates associated to a true photon with true E_{T} greater than the same threshold used to compute N_γ^{truth}.

The identification efficiency with respect to the reconstructed true candidates is defined as:

$$\varepsilon_{\gamma,\text{wrt reco}} = \frac{N_\gamma^{\text{truth, pass cut}}}{N_\gamma^{\text{truth, reco}}}\,, \tag{2.2}$$

where $N_\gamma^{\text{truth, reco}}$ is the total number of reconstructed photons that are associated to true photons with true E_{T} greater than the same threshold used to compute $N_\gamma^{\text{truth, pass cut}}$. The photon identification efficiency with respect to the generated true photon can thus be rewritten as

$$\varepsilon_\gamma = \varepsilon_{\gamma,\text{wrt reco}} \frac{N_\gamma^{\text{truth, reco}}}{N_\gamma^{\text{truth}}}\,, \tag{2.3}$$

where the second factor in the right-hand side of the equation is the photon reconstruction efficiency illustrated in Fig. 1.9.

The jet rejection, defined as the ratio between the total number of generated hadronic jets (at truth-particle level) and the number of jets reconstructed as photons and passing the photon identification criteria,

$$R_j = \frac{N_{\text{jets}}}{N_\gamma^{\text{fake,passcut}}}, \tag{2.4}$$

is computed from the fake photon yield $N_\gamma^{\text{fake,passcut}}$ in the jet filtered jet sample and the number of generator-level true jets, in the following way:

$$R_j = \frac{\frac{N_{\text{jets}}}{N_2}\frac{N_1}{\varepsilon_{\text{filter}}}}{N_\gamma^{\text{fake, pass cut}}} = \frac{N_{\text{jets}}}{N_2}\frac{N_1/\varepsilon_{\text{filter}}}{N_\gamma^{\text{fake, reco}}}\frac{1}{\varepsilon_{j,\text{wrt reco}}} \tag{2.5}$$

where

- $N_\gamma^{\text{fake, reco}}$ ($N_\gamma^{\text{fake, pass cut}}$) is the number of fake photons reconstructed (selected) in the filtered sample,
- N_1 is the number of events simulated in the filtered sample,
- N_{jets} is the total number of truth-particle hadronic jets (obtained by summing particle four-momenta within a cone size $\Delta R = 0.4$) in the unfiltered sample,
- N_2 is the size of the unfiltered sample,
- $\varepsilon_{\text{filter}}$ ($\approx 7\%$) is the efficiency of the generator filter applied before the simulation of the filtered sample.

$\frac{N_{\text{jets}}}{N_2}$ is thus the jet multiplicity in the generated events (from the high-statistics unfiltered sample), $N_1/\varepsilon_{\text{filter}}$ is the number of generated events that have been filtered before being passed through full simulation, $\frac{N_\gamma^{\text{fake, reco}}}{N_1/\varepsilon_{\text{filter}}}$ is the number of reconstructed fake photons per generated event of the filtered sample, and $\varepsilon_{j,\text{wrt reco}} = \frac{N_\gamma^{\text{fake, pass cut}}}{N_\gamma^{\text{fake, reco}}}$ is the efficiency of the photon identification criteria with respect to reconstructed fake photons (also referred to as "fake rate").

In the previous equations, the quantities that are directly affected by a redefinition of the identification criteria are $\varepsilon_{\gamma,\text{wrt reco}}$ and $\varepsilon_{j,\text{wrt reco}}$.

2.1.5 Optimization of the photon identification criteria

A first investigation of the photon trigger and offline identification criteria and of the identification efficiency for different signal samples using the default ATLAS selection at the beginning of 2009 shows that several of the requirements listed in Sec. 2.1.2 are not satisfied by the criteria used in Ref. [98]:

- the average efficiency for $E_{\mathrm{T}} > 20$ GeV is about 82% for photons from $H \to \gamma\gamma$ but only 70% for photons from SM γ-jet and $\gamma\gamma$ events, due to their softer transverse momentum spectra and the E_{T}-dependence of the efficiency.
- the efficiency is significantly different between unconverted and converted photons: in the case of $H \to \gamma\gamma$, the two efficiencies are respectively 87% and 73%.
- the selection criteria are not consistent with the photon trigger ones, for the following reasons:
  1. for consistency with the offline electron cut-based identification algorithm, the trigger implementation was changed in order to use thresholds tuned in seven pseudorapidity intervals differing from the six ones in which the offline tight photon criteria had been previously tuned.
  2. as the normalized total hadronic energy R_{had} was found to be more effective in discriminating hadronic showers in the region $0.8 < |\eta| < 1.37$ than $R_{\mathrm{had_1}}$ (see Figure 2.1), the trigger (again, for consistency with the offline electron cut-based identification algorithm) used R_{had} instead of $R_{\mathrm{had_1}}$ in that region.
- different cut thresholds has been chosen in eight E_{T} intervals.

To achieve the goals listed in Sec. 2.1.2, the tight photon identification criteria have thus been retuned. The optimization is performed separately for photon candidates reconstructed either as converted or unconverted ones, to take into account the different shower-shape distributions and to achieve similar efficiencies. Within the fiducial $|\eta|$ region, the cut values are determined independently in the same seven $|\eta|$ intervals used for the definition of the photon trigger identification criteria; a total of fourteen independent optimizations is thus performed. The cut thresholds are chosen in order to provide an identification efficiency close to 85% with respect to the initial collection of reconstructed candidates for both unconverted and converted photons with $E_{\mathrm{T}} > 20$ GeV, and to minimize the corresponding fake rate. For consistency with the photon selection at trigger level, R_{had} is used instead of $R_{\mathrm{had_1}}$ for $0.8 < |\eta| < 1.37$. Moreover, the variable E_{ratio} replaces $R_{\mathrm{max\,1}}$, which was found to be less powerful in discriminating isolated leading π^0's, as shown in Fig. 2.5. The loose selection – consisting of the photon trigger identification requirements, applied to the offline values of the DVs – is applied as a prerequisite in order to ensure that the tight criteria are not looser than the trigger ones.

The optimizations are performed using the "Toolkit for MultiVariate Analysis" (TMVA) [114, 115], a ROOT [116]–based framework for the tuning and evaluation of several multivariate techniques for the classification of events in terms of two categories, signal and background, based on the per-event values of some discriminating variables. The package implements several classifiers of different complexity; for the studies presented here, we rely on TMVA's multi-dimensional rectangular cut optimization, based on a genetic algorithm. For each value ε of the signal efficiency $\varepsilon_{\gamma,\mathrm{wrt\,reco}}$ between 0 and 1 in steps of 0.01, TMVA performs a numerical search of the values of the cuts on the discriminating variables that minimize the background efficiency $\varepsilon_{j,\mathrm{wrt\,reco}}$ while keeping the signal efficiency constant and equal to ε. Multiple sets of cut values on the discriminating variables are explored in parallel by the genetic algorithm, in order not to neglect the significant correlations among the discriminating variables. The signal and background samples are split in two halves, one used for *training*, *i.e.* to find the optimal values of the cuts, and one used for *testing* the results, *i.e.* to assess on independent samples the properties of the multivariate classifier output in terms of signal efficiency, background efficiency, signal-to-background separation, and to check for "overtraining" (whether the MVA learned statistical fluctuations from the training sample, which are not there or different in an independent test sample). For each photon conversion category and pseudorapidity interval in which the optimization is performed, the TMVA output consists of a curve of background vs signal efficiency, similar to those illustrated (for a simple selection based on a single DV) in Fig. 2.5, and for each point of the curve, the corresponding set

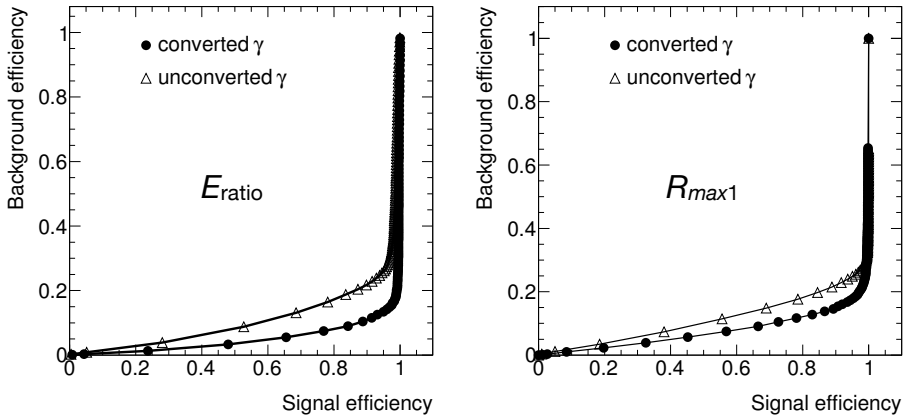

Figure 2.5: True (signal) versus fake (background) efficiency of a selection based only on the variable E_{ratio} (left) or the variable $R_{\mathrm{max\,1}}$ (right), for photon candidates reconstructed in $|\eta| < 0.6$ and with 25 GeV $< E_{\mathrm{T}} <$ 40 GeV before any selection, separately plotted for unconverted (open triangles) and converted (full circles) candidates.

of thresholds can be retrieved. In this study, for each optimization, the thresholds for the point of the curve closest to the point yielding $\varepsilon_{\gamma,\mathrm{wrt\,reco}} = 85\%$ and $R_j = 5000$ are chosen.

2.1.6 Results

Efficiencies and rejections of the reoptimized tight selection are computed for photon candidates in the fiducial volume in $0 < |\eta| < 1.37$ and $1.52 < |\eta| < 2.37$, separately for converted and unconverted candidates. For reference, also the efficiency and rejection of the loose photon identification requirements are quoted.

Table 2.1 summarizes the average photon identification efficiencies with respect to the reconstructed candidates, as defined in Eq. (2.2), computed for true prompt photons in the sample of γ-jet events having true E_{T} greater than 20, 25 or 40 GeV. For both unconverted and converted signal photons the tight selection efficiency with respect to the reconstructed candidates with $E_{\mathrm{T}} > 20$ GeV is about 85%, as initially foreseen.

Table 2.1: Expected photon identification efficiencies for loose and tight selections with respect to the reconstructed candidates.

		loose selection efficiency (%)			*tight* selection efficiency (%)		
		all	unconverted	converted	all	unconverted	converted
γ/j	$E_{\mathrm{T}} > 20$ GeV	98.22 ± 0.01	98.28 ± 0.01	98.11 ± 0.01	85.28 ± 0.02	85.47 ± 0.01	84.97 ± 0.01
	$E_{\mathrm{T}} > 25$ GeV	98.60 ± 0.01	98.54 ± 0.01	98.69 ± 0.01	87.65 ± 0.02	87.57 ± 0.03	87.79 ± 0.03
	$E_{\mathrm{T}} > 40$ GeV	99.00 ± 0.01	98.81 ± 0.01	99.29 ± 0.01	91.64 ± 0.03	91.15 ± 0.03	92.46 ± 0.04

Table 2.2 summarizes the average total photon selection efficiencies (including the reconstruction efficiency), as defined in Eq. (2.1), for true photons either from γ-jet events or from $H \to \gamma\gamma$ decays ($m_H = 120$ GeV) and having true E_{T} greater than 20, 25 or 40 GeV, and for true photons from $G \to \gamma\gamma$ decays ($m_G = 500$ GeV) having true E_{T} greater than 100 GeV.

Figure 2.6 shows the total expected selection efficiencies for loose and tight selections, for true prompt photons in a sample of γ-jet events in the region $0 < |\eta| < 1.37$ and $1.52 < |\eta| < 2.37$ having true $E_{\mathrm{T}} > 20$ GeV, as a function of the transverse energy.

Table 2.3 summarizes the average jet rejections, computed according to Eq. (2.5), of the loose and tight selections for all, quark–initiated and gluon–initiated jets having true E_{T} greater than 20, 25 or 40 GeV.

The loose and the tight selections documented here and in Ref. [10] have been the default photon identification algorithms in the ATLAS offline reconstruction and analysis software since the beginning of 2010 and have been used for most of the analyses of 2010 and 2011 data, with

Table 2.2: Expected total photon selection efficiencies (reconstruction + identification) for loose and tight selections.

		loose selection efficiency (%)			*tight* selection efficiency (%)		
		all	unconverted	converted	all	unconverted	converted
γ/j	$E_{\mathrm{T}} > 20$ GeV	95.45 ± 0.01	97.80 ± 0.01	91.73 ± 0.01	82.88 ± 0.02	85.04 ± 0.03	79.44 ± 0.04
	$E_{\mathrm{T}} > 25$ GeV	95.96 ± 0.01	98.08 ± 0.01	92.58 ± 0.03	85.31 ± 0.02	87.16 ± 0.03	82.35 ± 0.04
	$E_{\mathrm{T}} > 40$ GeV	96.37 ± 0.02	98.40 ± 0.02	93.17 ± 0.04	89.21 ± 0.03	90.76 ± 0.04	86.76 ± 0.05
$H \to \gamma\gamma$	$E_{\mathrm{T}} > 20$ GeV	96.15 ± 0.05	97.93 ± 0.04	93.00 ± 0.10	88.45 ± 0.08	89.81 ± 0.09	86.05 ± 0.14
	$E_{\mathrm{T}} > 25$ GeV	96.25 ± 0.05	97.99 ± 0.04	93.16 ± 0.10	88.92 ± 0.08	90.22 ± 0.09	86.61 ± 0.14
	$E_{\mathrm{T}} > 40$ GeV	96.46 ± 0.05	98.16 ± 0.04	93.42 ± 0.11	90.06 ± 0.08	91.27 ± 0.09	87.92 ± 0.14
$G \to \gamma\gamma$	$E_{\mathrm{T}} > 100$ GeV	95.91 ± 0.01	97.70 ± 0.01	92.55 ± 0.01	90.89 ± 0.01	91.73 ± 0.01	89.29 ± 0.01

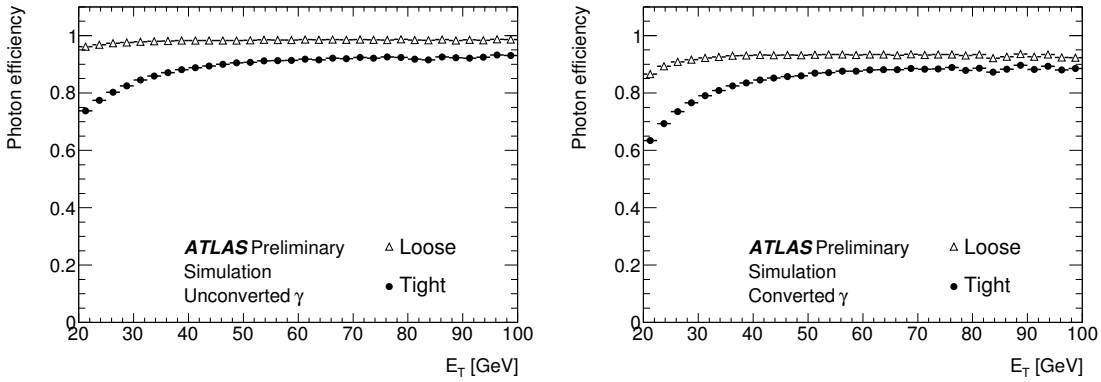

Figure 2.6: Expected total photon efficiency (reconstruction + identification) vs true E_{T} for loose and tight selection criteria and for unconverted (left) and converted (right) photons in the medium E_{T} range.

Table 2.3: Expected jet background rejections for loose and tight selections.

	loose rejection			tight rejection		
	all jets	quark jets	gluon jets	all jets	quark jets	gluon jets
$E_{\mathrm{T}} > 20$ GeV	898 ± 4	323 ± 2	2224 ± 16	4780 ± 43	1626 ± 17	13688 ± 238
$E_{\mathrm{T}} > 25$ GeV	1030 ± 6	365 ± 3	2674 ± 30	5288 ± 71	1743 ± 27	16765 ± 458
$E_{\mathrm{T}} > 40$ GeV	944 ± 13	368 ± 6	2211 ± 55	5098 ± 165	1675 ± 60	20110 ± 1504

some minor changes introduced later – when the first data started to be available – to reduce the systematic effects associated with the differences observed between the distributions of the calorimetric variables in data and in the simulation. Before the 2012 data-taking, based on the experience gained with data in 2011 and the significantly improved data/MC agreement obtained using photon control samples described in the next section, a full reoptimization of the photon identification criteria has been performed by other ATLAS collaborators.

2.2 Measurement of the photon identification efficiency with 7 and 8 TeV data

As mentioned at the beginning of this chapter, the precise determination of the photon identification efficiency is an important ingredient for accurate cross section measurements of processes involving photons in the final state. Unlike electrons, for which the $J/\psi \to ee$, $W \to e\nu$ and $Z \to ee$ decays provide clean control samples of electrons that can be used to measure *in situ* the electron identification efficiency [83], at the LHC there are no physics processes that produce high-statistics and clean samples of prompt photons over a large E_{T} range. The purest photon control sample is provided by the radiative $Z \to \ell\ell\gamma$ ($\ell = e, \mu$) decays, but the small value of their production cross section and their kinematics, characterized by a steeply-falling photon E_{T} spectrum, limit the energy range over which the efficiency $\varepsilon_{\mathrm{ID}}$ [2] of the photon identification criteria can be determined with sufficient accuracy. For these reasons, the first SM photon cross section measurements by ATLAS relied on MC-based estimates of $\varepsilon_{\mathrm{ID}}$, obtained after correcting the simulated values of the electromagnetic shower-shape variables by the average differences between their data and MC distributions in photon-enriched samples. The uncertainty on these MC-based $\varepsilon_{\mathrm{ID}}$ values were mainly associated to the correction technique, accounting for the imperfect knowledge of the material upstream of the electromagnetic calorimeter, the uncertainty on the photon candidate purity in the sample used to derive the corrections, and the accuracy of the data/MC discrepancy parametrizations used to correct the MC. This approach led to a rather large uncertainty on the photon identification efficiency.

With more data available, however, it has become possible to determine $\varepsilon_{\mathrm{ID}}$ *in situ*, thus reducing the corresponding systematic uncertainties on the measured photon production cross sections. Measurements of $\varepsilon_{\mathrm{ID}}$ have thus been performed with the full 2011 and 2012 pp collision data. Three different data-driven techniques have been investigated: selecting photons from radiative decays of the Z boson [117], extrapolating photon properties from electrons and positrons from Z decays by exploiting the similarity of electron and photon electromagnetic showers [118], and implementing a technique to determine the fraction of background present in samples of isolated photon candidates either passing or failing the identification requirements ("matrix" approach) [119]. The three techniques allow the measurement of $\varepsilon_{\mathrm{ID}}$ in complementary E_{T} regions (low, medium and high-E_{T}, respectively). With Kun Liu, a student that started in September 2011 his Ph.D. studies in physics, in co-tutorship between UPMC (Paris, France) and USTC (Hefei, Chine) under the direction of prof. Yanwen Liu and mine, I worked on the first and the third methods; in the following I will describe them briefly, together with the final results. More details are provided in Refs. [11, 12, 117, 119, 120].

The efficiency of the tight identification requirements is measured with respect to true photons that are reconstructed and pass a calorimeter-isolation requirement. The normalization of the efficiency with respect to isolated photons reduces to a negligible level the differences of the identification efficiency between photons from decays of heavy resonances (like $H \to \gamma\gamma$ or $Z \to \ell\ell\gamma$), which are more isolated and characterized by narrower shower shapes, and photons from QCD scattering events (selected with the matrix method technique), where in addition to the isolated component from the hard-scattering events (like $qg \to q\gamma$) there is a less-isolated component, from parton fragmentation, characterized by broader shower shapes. As a matter of fact, most if not all the ATLAS photon analyses impose a photon isolation requirement from the beginning, in order to improve the purity of the selected photon sample, and in several cases – as in the measurements of the SM prompt photon cross sections – this is even part of the particle-level definition of the cross

2. denoted as $\varepsilon_{\gamma,\mathrm{wrt\,reco}}$ in the previous section

section. All these analyses therefore need as input the identification efficiency for isolated photons. For 7 (8) TeV data, the efficiency is measured for photons having a calorimeter isolation, computed using all the calorimeter cells (belonging to topological clusters) in a cone of radius $R = 0.4$, less than 5 (4) GeV. These requirement have been widely used in most of the ATLAS photon analyses based on 2011 and 2012 data. The efficiency is measured in several bins of transverse energy and in the following four intervals of the absolute value of the pseudorapidity:

$$[0, 0.6) \quad , \quad [0.6, 1.37) \quad , \quad [1.52, 1.81) \quad , \quad [1.81, 2.37)$$

The first (last) two intervals correspond to photons in the barrel (end-cap) of the ECAL; in the interval $[0.6, 1.37)$ the material upstream of the calorimeter is significantly larger than for $|\eta| < 0.6$. In the end-cap, the two intervals correspond to photons passing or not through the presampler.

2.2.1 Previous studies

In the ATLAS performance projections performed before the start of the data-taking [121], relative uncertainties on the identification efficiency of 1.0%, 0.5% and 0.2% have been assumed for three values of the integrated luminosity, 0.1, 1 and 10 fb^{-1}, respectively, at $\sqrt{s} = 14$ TeV, based on the expected size of the electron control sample from $Z \to ee$ decays and on the hypothesis that the efficiency of the photon identification criteria could be estimated directly by applying those criteria to the shower shapes of electron candidates. However, these optimistic estimates have been challenged when ATLAS started to take data in 2010 at $\sqrt{s} = 7$ TeV: the smaller $pp \to Z$ cross section at the lower CM energy yielded a smaller control samples of electrons than expected, and extrapolating from electron to photon shower shapes turned out to be not as straightforward as anticipated, particularly for unconverted photons. During 2010 and the following years, a constant effort aiming to determine with ever increasing accuracy the photon identification efficiency has thus been put in place.

The first determination of the photon identification efficiency used in an ATLAS publication based on data has been performed for the measurement of the inclusive isolated prompt photon production cross section at $\sqrt{s} = 7$ TeV, for photons with $|\eta| < 1.81$, using the first ≈ 1 pb^{-1} of ATLAS data collected in 2010 [14]. It is based on a "corrected" prompt photon simulation, where the shower-shape variables are shifted by correction factors computed as the differences between the means of the distributions of each discriminating variable in data and in a MC sample containing all the main QCD signal and background processes, after applying the tight identification criteria [3]. The typical size of the correction factors is 10% of the RMS of the distribution of the corresponding variable in data, with a maximum of 50% of the RMS for the variable (R_η) where the simulation is in worse agreement with the data [4]. The correction to the MC efficiency ranges between -5% and zero, with a typical value of $\approx -3\%$. The systematic uncertainties on the photon identification efficiency are computed using alternative samples, differing for the amount of material upstream of the calorimeter, the pile-up conditions, the event generators used (with different underlying event and hadronization models), the fraction of fragmentation photons, and the probability to reconstruct a converted photon as unconverted, Conservative estimates of the uncertainties of the previous quantities are used. The dominant source of uncertainty is the amount of material upstream of the calorimeter (absolute uncertainties on $\varepsilon_{\rm ID}$ ranging between 1% and 8%, larger at low $E_{\rm T}$). An additional systematic uncertainty due to the shower-shape correction procedure is computed from a closure test based on two simulated prompt photon samples, one using the nominal detector geometry and one using a model, containing an additional 10% of material in the inactive volumes of the inner detector and 10% of a radiation length in front of the electromagnetic calorimeter, that is estimated to represent a conservative upper limit of the additional detector material that is not accounted for by the nominal simulation. The differences between the averages of the shower shapes distributions in the two samples are computed, and used to correct the shower shapes of the nominal simulation; the photon efficiency from the nominal simulation is recomputed after applying these corrections, and compared to the efficiency obtained from the alternative simulation.

3. Later, this technique has been refined by computing the correction factors as the shift values that minimize a χ^2 computed from the data and MC histogrammed distributions

4. the agreement between data and Monte Carlo has been significantly improved in 2011 and 2012, reducing the size of the data-MC corrections by about one half

Their difference ranges from 3% at $E_{\mathrm{T}} \approx 20$ GeV to less than 1% at $E_{\mathrm{T}} \approx 80$ GeV, corresponding to relative uncertainties of 5% and 1% respectively. Overall, the absolute (relative) uncertainty on the photon efficiency obtained with this method is around 8% (12%) for $E_{\mathrm{T}} = 15$ GeV and 3% (3%) for $E_{\mathrm{T}} = 100$ GeV. The efficiency is a few % lower than anticipated from the studies described in Sec. 2.1, due to the mismodeling of the photon shower shapes in the simulated samples used in those studies. A similar method has been used to compute the efficiency of the photon identification criteria for the measurement of the SM di-photon production cross section using all the 7 TeV data collected in 2010 (37 pb^{-1}). The efficiency, for events with two photons with $E_{\mathrm{T}} > 16$ GeV, increases as a function of the di-photon invariant mass $m_{\gamma\gamma}$ between 55% and 75%, with a relative uncertainty around 13-17% [17]. The measured photon identification efficiency for single and di-photon events is shown in Fig. 2.7. A systematic uncertainty of 11% on the $H \to \gamma\gamma$ selection efficiency from the MC-based estimate of the photon identication efficiency has also been used in the first ATLAS $H \to \gamma\gamma$ public results [21, 24].

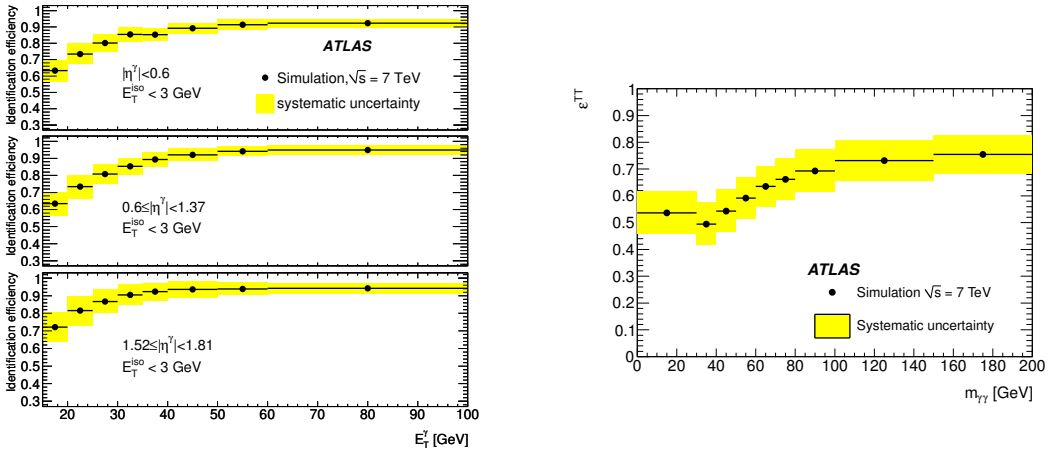

Figure 2.7: Left: efficiency of the tight identification criteria as a function of the reconstructed photon transverse energy for isolated prompt photons, estimated from a simulation of SM photon-jet events where the shower-shape variables have been corrected for the average data-MC differences [14]. Right: efficiency of the tight identification criteria as a function of the di-photon invariant mass, from a corrected simulation of SM isolated-photon pair events where both photons have $E_{\mathrm{T}} > 16$ GeV. [17]. In both figures, systematic uncertainties are included. The curves correspond to the photon identification criteria used for the analysis of the $\sqrt{s} = 7$ TeV ATLAS data collected in 2010.

2.2.2 Measurement of the photon identification efficiency in 2011 data using Z radiative decays

A sample of isolated photons from radiative Z decays ($Z \to \ell\ell\gamma$) can be obtained in data, with small background contamination, by applying a selection only to identification variables characterizing the two leptons from the Z decay and to the kinematics of the di-lepton and the $\ell\ell\gamma$ three-body systems. The DVs from such a photon sample are thus unbiased, and the sample can be used to extract $\varepsilon_{\mathrm{ID}}$ and to perform data/MC shower shape comparisons. Due to the rapid decrease of the cross section with increasing photon E_{T}, the measurement is restricted to the $15 < E_{\mathrm{T}} < 50$ GeV range, as the statistical uncertainty on $\varepsilon_{\mathrm{ID}}$ becomes too large at higher energies. This, however, provides a data-driven insight in the region where the MC $\varepsilon_{\mathrm{ID}}$ prediction is most sensitive to several sources of systematic uncertainty, such as imperfect modeling of the passive material in front of the ECAL. Consistent results are obtained using the two channels, $Z \to \mu\mu\gamma$ and $Z \to ee\gamma$; they are thus be combined together into a single measurement.

Events are collected using the lowest-threshold, unprescaled single-lepton or di-lepton triggers, with p_{T} thresholds around 20 GeV and 12 GeV, respectively. Only events where all the subde-

tectors are operative and with good data quality are retained: they correspond to an integrated luminosity of about 4.9 fb^{-1} at $\sqrt{s} = 7$ TeV. Events are further selected by requiring two same-flavor, oppositely-charged leptons, with transverse momenta above 15 GeV and pseudorapidities within a fiducial region of either the electromagnetic calorimeter (for electrons: $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$) or the muon spectrometer and the inner detector (for muons: $|\eta| < 2.4$). The leptons are required to be isolated, based either on energy deposited in the calorimeter near the electron cluster (less than 5 GeV within a cone of radius 0.4 around the electron) or the scalar sum of the transverse momenta of the tracks with $p_{\mathrm{T}} > 500$ MeV in a cone of radius 0.2 around the muon (less than 10% of the muon p_{T}). Quality criteria, aimed at suppressing fake candidates from jets of hadrons, are applied to the tracking information and calorimeter shower shape of electrons and to the number of hits of the muons in the pixel (≥ 1) and silicon strip (≥ 6) detectors. To suppress pile-up background, both lepton tracks reconstructed in the ID are required to have a longitudinal impact parameter, with respect to the hard-scattering primary vertex, smaller than 10 mm, and transverse impact parameter significance, $|d_0|/\sigma_{d_0}$, less than 10.

As the cross section for Z+jets events is about three orders of magnitude higher than that for $Z + \gamma$ events, and a non-negligible fraction, $\mathcal{O}(1\%)$, of jets with a significant electromagnetic fraction (from neutral mesons decaying to collimated photon pairs) can be reconstructed as a photon candidate, a large jet background is present in the initial-state-radiation (ISR) $\ell\ell\gamma$ sample. In order to minimize the impact of such background, a kinematic selection that suppresses the contribution of ISR events compared to final-state-radiation (FSR) ones, in which the photon is radiated off one lepton, is exploited. Figure 2.8 shows the two-dimensional distribution of the three-body $\ell\ell\gamma$ and two-body $\ell\ell$ invariant masses in data events passing the criteria listed above, separately for the electron and muon channels. FSR events have a three-body invariant mass $m_{\ell\ell\gamma}$ near the Z boson mass (m_Z) and a di-lepton invariant mass $m_{\ell\ell} < m_Z$, whereas Z+jets and ISR events lie in a distinct region of the phase-space, characterized by $m_{\ell\ell} \approx m_Z$, $m_{\ell\ell\gamma} > m_Z$: we thus require 80 GeV $< m_{\ell\ell\gamma} <$ 96 GeV and 40 GeV $< m_{\ell\ell} <$ 83 GeV.

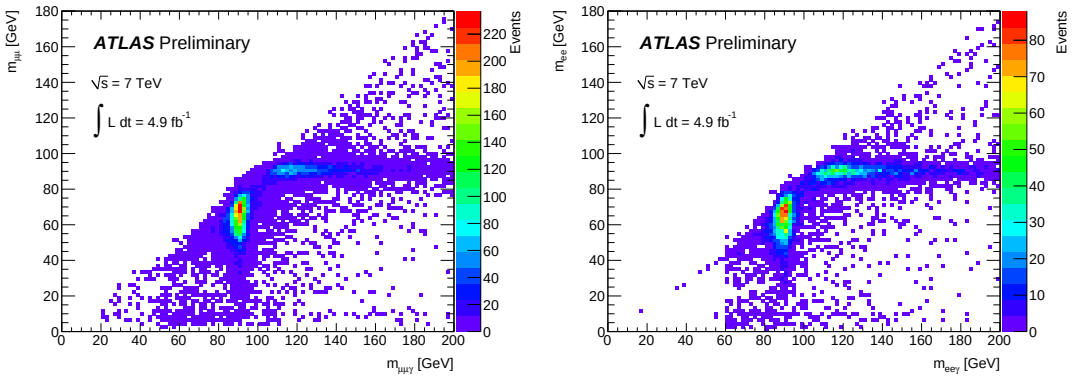

Figure 2.8: Distribution of the invariant masses $m_{\ell\ell\gamma}$ vs $m_{\ell\ell}$ from photon candidates in the full 2011 data set after event selection, for $\mu\mu\gamma$ (left) and $ee\gamma$ (right) events.

Photons emitted in FSR events tend to be collinear with the radiating lepton, and the cross section is strongly peaked for $\Delta R_{\mathrm{min}} \to 0$, where ΔR_{min} is the minimum $\{\eta, \phi\}$ separation between the photon and the two leptons. If one of the two leptons is close enough to the photon, the energy deposited by the lepton in the calorimeter can alter the shower-shape variables of the photon, biasing them towards lower $\varepsilon_{\mathrm{ID}}$: for this reason, we require $\Delta R_{\mathrm{min}} > 0.2$ for the muon channel and $\Delta R_{\mathrm{min}} > 0.4$ for the electron channel. In the simulation it is verified that the efficiency of the identification criteria for ΔR_{min} larger than those values is compatible within statistical uncertainties to that obtained from simulated QCD prompt photons.

After the selection, 5995 (2705) events containing unconverted (converted) photon candidates, or "probes", remain in the muon channel. In the electron channel, 2481 (1057) events containing unconverted (converted) photon probes are selected. The efficiency of the photon tight identifica-

tion criteria is evaluated as the fraction of the probes that pass the tight identification criteria:

$$\varepsilon_{\mathrm{ID}} = \frac{N_{\mathrm{probes,tight}}}{N_{\mathrm{probes}}}. \tag{2.6}$$

The uncertainty on $\varepsilon_{\mathrm{ID}}$ is dominated by the statistical component, which is maximum ($\pm 5\%$) in the high photon E_{T} (> 30 GeV) and high $|\eta|$ region. A small systematic uncertainty originates from the presence, in the selected photon sample, of (mostly Z+jet) background that is neglected in the nominal result. Maximum-likelihood fits to the $m_{\ell\ell\gamma}$ invariant mass distribution in data are used to estimate the residual background contamination, before and after applying the tight photon identification criteria, and its impact on the efficiency measurement. The data distribution is fitted with the function $N_S f_S(m_{\ell\ell\gamma}) + N_B f_B(m_{\ell\ell\gamma})$, where f_S and f_B are the signal and background $m_{\ell\ell\gamma}$ probability density functions (*pdf*s), obtained from MC simulation, and N_S and N_B are the (floating) signal and background yields. The fits are performed by relaxing the $m_{\ell\ell\gamma}$ requirement to 60 GeV $< m_{\ell\ell\gamma} <$ 140 GeV. Figure 2.9 shows that the fitted sum is in good agreement with

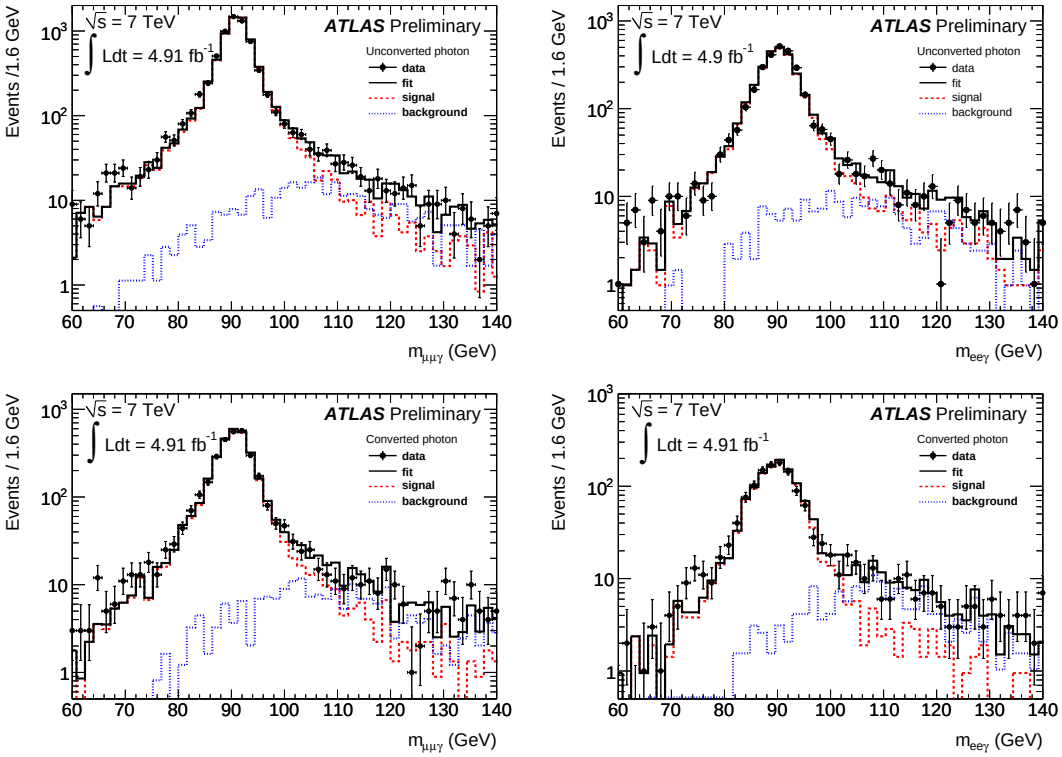

Figure 2.9: Invariant mass ($m_{\ell\ell\gamma}$) distribution of events selected in data after applying all the $Z \to \ell\ell\gamma$ selection criteria except that on $m_{\ell\ell\gamma}$ (left: $\ell = \mu$; right: $\ell = e$), for unconverted (top) and converted (bottom) photon candidates. The result of the fit to the data distributions with the sum of the signal and background invariant mass distributions, as obtained from MC simulation, is superimposed.

the data distribution. From the fits the average photon purity is estimated to be $(98.4 \pm 0.2)\%$ for the $\mu\mu\gamma$ channel for unconverted photons, and $(98.2 \pm 0.3)\%$ for converted photons; for $ee\gamma$ it is $(98.0 \pm 0.2)\%$ for unconverted photons and $(96.0 \pm 0.3)\%$ for converted photons, which suffer from $e \leftrightarrow \gamma$ misidentification. The measured change in the identification efficiency after subtracting the estimated background is at most 0.7% for unconverted photons and 1.6% for converted photons in the $Z \to \mu\mu\gamma$ channel. For $Z \to ee\gamma$, the difference is 0.8% for unconverted photons and 2.5% for converted photons. The number of events available in either the data sample or the MC sample does not allow this study to be carried out independently for each photon E_{T} and η region considered. Therefore, this overall difference is considered as the systematic uncertainty associated with the neglected background in each (E_{T}, η) bin.

2.2.3 Measurement of the photon identification efficiency in 2011 data with the matrix method

An inclusive sample of photon candidates is collected by selecting events passing single photon triggers and requiring the presence of at least one reconstructed photon candidate, with $E_{\mathrm{T}} > 20$ GeV, isolated in the calorimeter and matched to a photon trigger object. The collected sample is then divided in two subsamples, the sets P (F) of candidates passing (failing) the tight photon identification criteria. The observed numbers of photon candidates in the two subsamples, N_p and N_f respectively, can be expressed in terms of signal S (*i.e.* true photons) and background B (fakes) candidates that pass or fail those requirements, leading to two equations with four unknowns:

$$\begin{aligned} N_p &= S_p + B_p \\ N_f &= S_f + B_f \end{aligned} \tag{2.7}$$

Photons are discriminated from fake candidates on a statistical basis by means of their track isolation, defined as the number of ID tracks with $p_{\mathrm{T}} > 500$ MeV in a cone of $0.1 < \Delta R < 0.3$ around the direction of a photon candidate. A photon candidate is considered track-isolated if no track is found in the cone. If the track-isolation efficiencies for signal and background candidates passing (ε_p^S, ε_p^B) or failing (ε_f^S, ε_f^B) the identification criteria are known, the photon purities P_p and P_f in the P and F samples, as well as the photon identification efficiency $\varepsilon_{\mathrm{ID}} = \frac{S_p}{S_p+S_f} = \frac{P_p N_p}{P_p N_p + P_f N_f}$, can be determined in the following way. The observed number of track-isolated photon candidates that either pass (N_p^I) or fail (N_f^I) the tight identification criteria yields two additional constraints on the four yields (S_p, B_p, S_f, B_f):

$$\begin{aligned} N_p^I &= \varepsilon_p^S S_p + \varepsilon_p^B B_p \\ N_f^I &= \varepsilon_f^S S_f + \varepsilon_f^B B_f \end{aligned} \tag{2.8}$$

We are thus left with a system of four equations in the four unknown yields, that can be solved explicitly. The signal purities P_p and P_f are

$$P_p = \frac{\varepsilon_p - \varepsilon_p^b}{\varepsilon_p^s - \varepsilon_p^b}\,, \tag{2.9}$$

$$P_f = \frac{\varepsilon_f - \varepsilon_f^b}{\varepsilon_f^s - \varepsilon_f^b}\,, \tag{2.10}$$

where $\varepsilon_{p,f} \equiv \frac{N_{p,f}^I}{N_{p,f}}$ are the fractions of tight or non-tight photon candidates in data that pass the track isolation criteria. Since the photon triggers apply some loose requirements on the electromagnetic shower shapes, they may slightly bias the ID efficiency measurement, especially at low transverse momentum. The measured efficiency is thus multiplied by a correction factor, equal to the ratio between the tight ID efficiency for all reconstructed photons and that for photons matching the trigger object that triggers the event, as obtained from a corrected simulation of prompt photon events:

$$\frac{\varepsilon_{\mathrm{ID}}}{\varepsilon_{\mathrm{ID,trig}}} = \frac{\frac{N_\gamma^{\mathrm{tight,reco}}}{N_\gamma^{\mathrm{reco}}}}{\frac{N_\gamma^{\mathrm{tight,reco,trig}}}{N_\gamma^{\mathrm{reco,trig}}}} \tag{2.11}$$

As an independent cross-check, the same factor can also be obtained, by simply rearranging the terms of the right-hand-side:

$$\frac{\varepsilon_{\mathrm{ID}}}{\varepsilon_{\mathrm{ID,trig}}} = \frac{\frac{N_\gamma^{\mathrm{reco,trig}}}{N_\gamma^{\mathrm{reco}}}}{\frac{N_\gamma^{\mathrm{tight,reco,trig}}}{N_\gamma^{\mathrm{tight,reco}}}}\,, \tag{2.12}$$

from the ratio of the trigger efficiency for tight photons and the trigger efficiency for all reconstructed photons, that can be measured directly in data using for instance photon probes from radiative $Z \to \ell\ell\gamma$ decays.

Measuring $\varepsilon_{\mathrm{ID}}$ in this way, by obtaining the signal component before and after tight cuts in the P and F subsamples, is referred to as the matrix method. It has the substantial advantage of providing a data-driven measurement of $\varepsilon_{\mathrm{ID}}$ over the whole E_{T} spectrum of the photon candidate data sample, particularly at high E_{T} where the other two methods are limited by the scarce statistics of the control sample.

The measurement presented here is based on the full ATLAS data set collected in 2011. Events are required to pass any of the inclusive photon triggers with E_{T} thresholds of 20, 40, 60 and 80 GeV. The use of several triggers with different thresholds has been necessary because lower-threshold triggers started to be prescaled as soon as the rapidly increasing peak luminosity reached a level where the photon trigger rate became too high. The integrated luminosities of each trigger range between 14 pb^{-1} for the 20 GeV trigger and 4.9 fb^{-1} for the (unprescaled) 80 GeV trigger. The signal track-isolation efficiency is obtained from a simulation of photon-jet events generated with PYTHIA (including both photon production in the hard-scattering as well as in radiation off partons from QCD $2 \to 2$ events). Its uncertainty is estimated by comparing the track-isolation efficiency in data and simulation for a pure sample of electrons selected from $Z \to ee$ decays with a tag-and-probe technique, after requiring the electrons to pass (or fail) the unconverted or converted photon identification criteria. The uncertainties in the signal track-isolation efficiency are found to be relatively small, less than $\pm 5\%$ for converted photons failing the tight criteria and $\pm 1\%$ in the rest of the cases. To estimate the uncertainty from the fraction of fragmentation photons we use simulated photon-jet events generated with an alternative program, SHERPA. The photon track isolation efficiency in SHERPA is in good agreement with that from PYTHIA once the effects from the different underlying event models are taken into account (by normalizing the efficiency with respect to the electron track isolation efficiency) and the effect on the final photon identification efficiency is estimated to be smaller than 1% and neglected. The background track-isolation efficiencies are estimated from a data sample enriched in fake photons, selected by reversing the tight selection on four shower shape variables computed from a few narrow cells of the first layer of the ECAL near the photon candidate, F_{side}, $w_{s\,3}$, ΔE, and E_{ratio}, which have small or negligible correlations with track isolation. This sample contains no background passing tight identification cuts by construction, thus in order to obtain ε_p^B a *relaxed* tight selection, consisting of those events which fail the cuts on the four DVs but pass the rest of the tight selection, is defined. Due to the very small correlation between the track isolation and these shower shape variables, the background track-isolation efficiency is similar for fake photons passing tight cuts or relaxed-tight criteria. This hypothesis is tested with di-jet MC simulated samples; the differences are included in the systematic uncertainties. Figure 2.10 shows as an example the true and fake

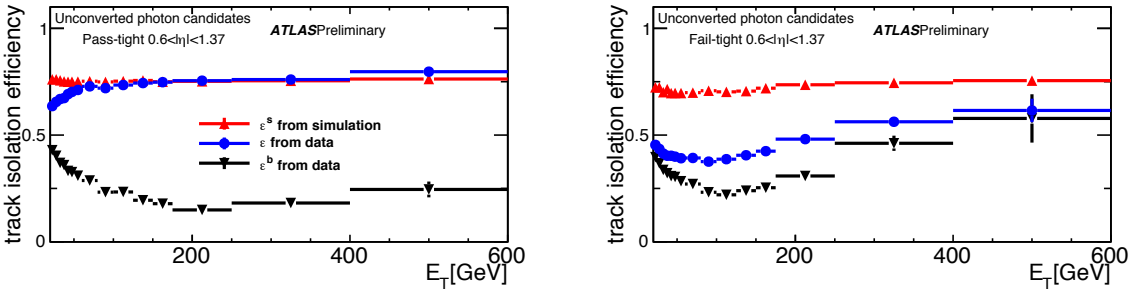

Figure 2.10: Track isolation efficiencies for unconverted signal and background photon candidates with $0.6 < |\eta| < 1.37$ passing (ε_p: left figure) or failing (ε_f: right figure) the tight criteria, as measured in data and from MC simulated samples. The red triangles show the signal track isolation efficiency as obtained from simulation, whereas the black triangles show that of background, as determined from data using the relaxed-tight photon control sample. The blue circles show the overall track isolation efficiency found in data (signal and background together). Only statistical uncertainties are shown.

photon track-isolation efficiencies $\varepsilon_p^{S,B}$ and $\varepsilon_f^{S,B}$ as a function of the photon candidate E_{T} for unconverted photon candidates reconstructed in the pseudorapidity interval $0.6 < |\eta| < 1.37$. The photon purity and the efficiency of the tight photon identification selection are deduced from such

curves. The typical correction to account for the trigger bias (Eq. (2.11)) is estimated from the simulation to be about 95% for $20 < E_{\mathrm{T}} < 30$ GeV and approches 1 for increasing E_{T}. It has been verified, using a control sample of radiative Z decays, that this correction is correctly estimated by the simulation within 2% for $E_T = 20$ GeV and within less than 1% for $E_T > 40$ GeV.

2.2.4 Final results (2011 data)

The $\varepsilon_{\mathrm{ID}}$ curves for reconstructed photons as a function of the transverse momentum in the four pseudorapidity ranges under study, obtained independently from the three data-driven methods, are found to be in good agreement in the overlapping E_{T} regions [11] and are thus combined. Figures 2.11 and 2.12 show the comparison between the combined data-driven $\varepsilon_{\mathrm{ID}}$ curves and the predictions from the simulation of photon-jet events (either before or after applying the shower-shape corrections), for unconverted and converted photons, respectively. Data and corrected simulation are in reasonable agreement within their uncertainties, while the nominal simulation over-estimates $\varepsilon_{\mathrm{ID}}$, since the electromagnetic showers from photons in the simulation are narrower than those in data. The disagreement between data and corrected MC is almost always within $\pm 5\%$, with the single exception of unconverted photons in the high-$|\eta|$ region, for which it can become as large

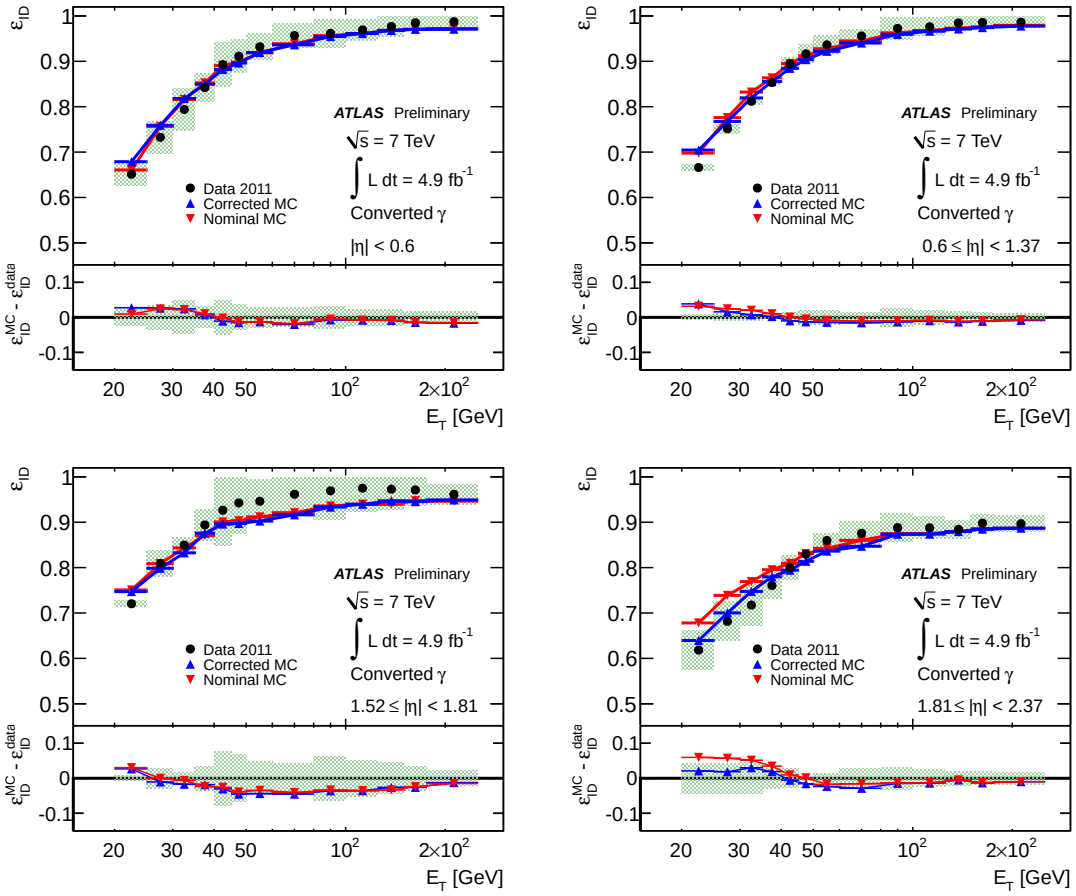

Figure 2.11: Comparison of the weighted mean of the data-driven measurements of converted $\varepsilon_{\mathrm{ID}}$ to the nominal and corrected MC predictions in the region 15 GeV $< E_{\mathrm{T}} <$ 300 GeV. The $\varepsilon_{\mathrm{ID}}$ curves are shown in four different η regions. The green uncertainty band corresponds to the addition in quadrature of the statistical and systematic uncertainties estimated for the combination of the data-driven methods. Only the statistical uncertainties are shown for the MC predictions. The bottom insets in each figure show the difference between the data-driven curve and nominal and corrected MC predictions.

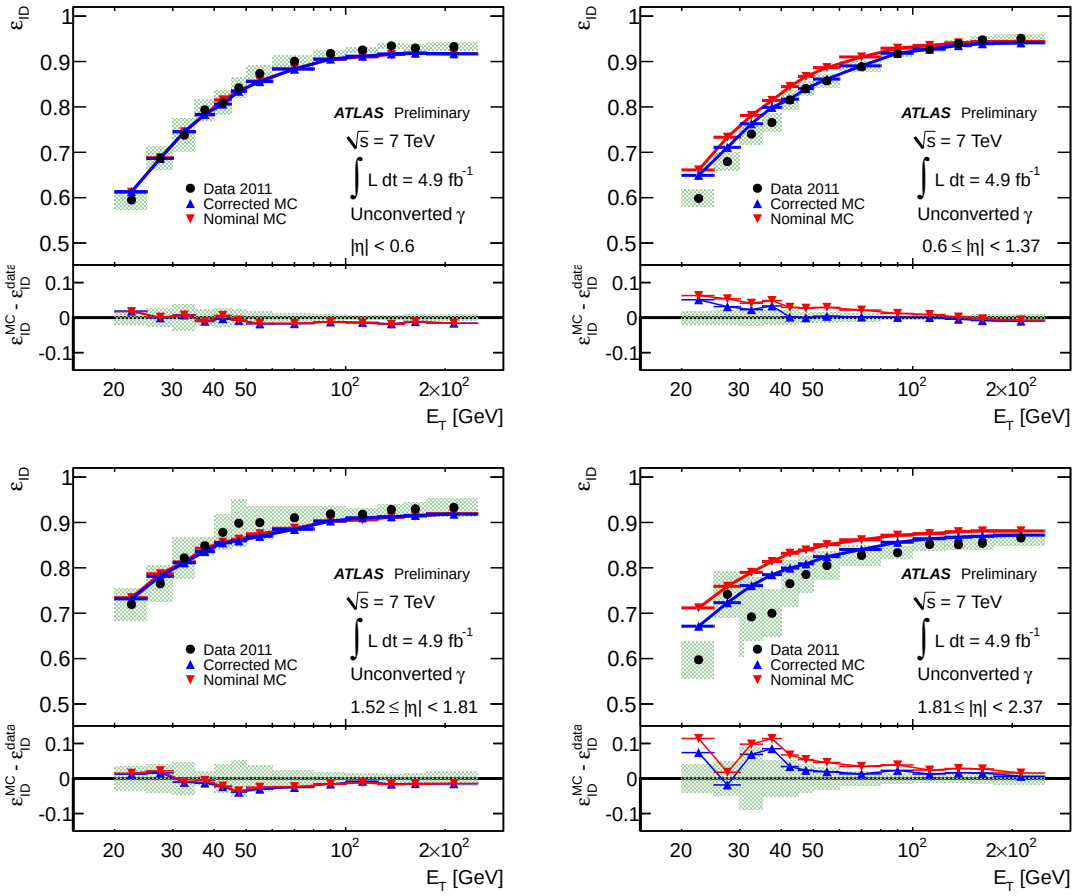

Figure 2.12: Comparison of weighted mean of the data-driven measurements of unconverted $\varepsilon_{\mathrm{ID}}$ to the nominal and corrected MC predictions in the region 15 GeV $< E_{\mathrm{T}} <$ 300 GeV. The $\varepsilon_{\mathrm{ID}}$ curves are shown in four different η regions. The green uncertainty band corresponds to the addition in quadrature of the statistical and systematic uncertainties estimated for the combination of the data-driven methods. Only the statistical uncertainties are shown for the MC predictions. The bottom insets in each figure show the difference between the data-driven curve and nominal and corrected MC predictions.

as ±8%. It reduces with increasing E_{T} and becomes negligible for $E_{\mathrm{T}} > 100$ GeV. Considering the difference between data and Monte Carlo efficiencies as the systematic uncertainty on $\varepsilon_{\mathrm{ID}}$ leads to a significant reduction (up to a factor of two) of the single and di-photon identification efficiency uncertainties described previously. For the latest cross-section measurement with photons at $\sqrt{s} = 7$ TeV, the bin-by-bin ratios between the data and corrected MC efficiency curves have been used to determine (E_{T}, η)-dependent scale factors that are applied to the simulation of the process under study ($H \to \gamma\gamma$, SM di-photons, ...). The systematic uncertainty due to photon identification is obtained by propagating the uncertainty on the scale factors themselves, and in particular the small uncertainty on the weighted average of the three data-driven methods. Additional minor contributions, estimated from the simulation, are included in the case of measurements using a different calorimeter-isolation requirement than the one used for measuring $\varepsilon_{\mathrm{ID}}$, or photon samples with a different contribution from parton-to-photon fragmentation compared to photon-jet events. This study allowed a significant reduction of the photon identification efficiency systematic uncertainty on the SM di-photon selection efficiency in 7 TeV data, to around 3.6% (relative) for two photons with $E_{\mathrm{T}} > 25$ and 22 GeV, respectively and invariant mass between 0 and 800 GeV, as shown in Fig. 2.13. For the $H \to \gamma\gamma$ search, since the photons have harder E_{T} spectra than in SM di-photon events and the uncertainty on $\varepsilon_{\mathrm{ID}}$ decreases with E_{T}, the uncertainty on the selection efficiency is even smaller, below 3%, which constitutes an improvement by a factor four compared to the previous uncertainty estimates based on the simulation.

Figure 2.13: Efficiency of the tight identification criteria as a function of the di-photon invariant mass, from a corrected simulation of SM isolated-photon pair events where the two photons have $E_{\mathrm{T}} > 25$ and 22 GeV, respectively [18]. Systematic uncertainties are included. The curve corresponds to the photon identification criteria used for the analysis of the $\sqrt{s} = 7$ TeV ATLAS data collected in 2011.

2.2.5 Photon identification efficiency with 2012 data - preliminary studies

Prior to taking data at $\sqrt{s} = 8$ TeV in 2012, the tight photon-identification criteria have been re-optimized. Moreover, the 2012 data-taking conditions are different (with much larger pile-up) from those of 2011, as well as the detector conditions (reconstruction algorithms, ..). For the analysis of the full 2012 data set, therefore, a new measurement of the photon identification efficiency is needed. To this purpose, in particular for the first public results of the $H \to \gamma\gamma$ and $H \to Z\gamma$ searches shown at the Moriond 2013 conferences [122, 28], with Kun Liu and Yanwen Liu I have performed a preliminary photon efficiency measurement at $\sqrt{s} = 8$ TeV with the Z radiative decay control sample [12]. The analysis strategy is identical to the one described in Sec. 2.2.2, with minor differences in some of the selection criteria in order to improve the efficiency of the selection and to take into account the different machine and detector conditions. Using the full 2012 data set

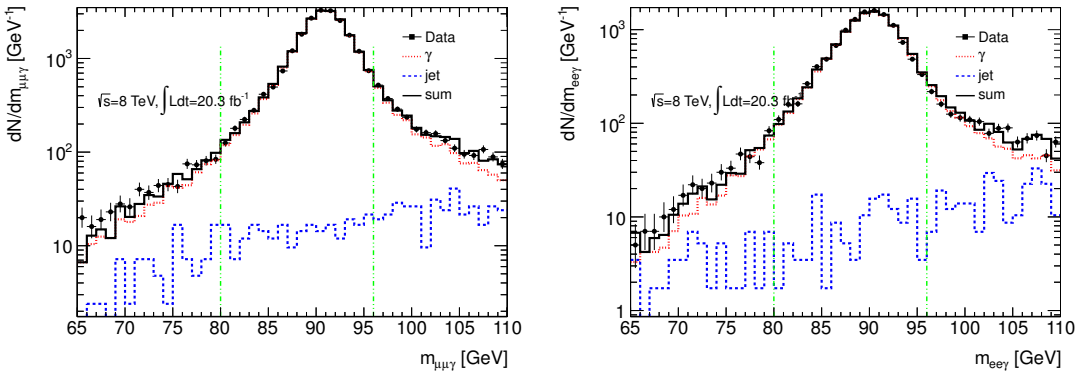

Figure 2.14: Invariant mass ($m_{\ell\ell\gamma}$) distribution of events selected in the full $\sqrt{s} = 8$ TeV data set after applying all the $Z \to \ell\ell\gamma$ selection criteria except that on $m_{\ell\ell\gamma}$ (left: $\ell = \mu$; right: $\ell = e$), for unconverted (top) and converted (bottom) photon candidates. The result of the fit to the data distributions with the sum of the signal and background invariant mass distributions, as obtained from MC simulation, is superimposed. Photons are required to be isolated in the calorimeter and have $E_{\mathrm{T}} > 20$ GeV.

(≈ 20.3 fb^{-1}), we select about 44026 photon probes with $E_{\mathrm{T}} > 10$ GeV in the electron channel and 73823 in the muon channel, with an estimated photon purity close to 90% for $10 < E_{\mathrm{T}} < 15$ GeV and around 98% for $E_{\mathrm{T}} > 15$ GeV, as determined from maximum-likelihood fits to the data $m_{\ell\ell\gamma}$ distribution (exemplified in Fig. 2.14). The resulting efficiencies are illustrated in Figs. 2.15 and 2.16. A good agreement (typically within 1.5% absolute for photons with $E_{\mathrm{T}} > 35$ GeV) is observed between the data-driven efficiency estimates and the expectations from the corrected simulation; from the differences between the two curves, a systematic uncertainty as small as 2.4% on the $H \to \gamma\gamma$ selection efficiency has been estimated. The results from the other two data-driven methods are still being computed; once they will be available, the uncertainty will be even smaller, and approaching the design value.

2.3 Optimization of the photon triggers for the 8 TeV run

During the winter 2011-2012 shutdown between the $\sqrt{s} = 7$ TeV and 8 TeV data-taking periods, the ATLAS trigger menu was completely revisited to cope with the expected running conditions for 2012, characterized by a larger center-of-mass energy and thus higher cross sections, a higher peak luminosity, anticipated to be around 7×10^{33} cm^{-2}s^{-1} throughout the whole 2012, and larger pile-up (with $20 - 30$ minimum-bias interactions per bunch crossing). The rates of the main triggers used in 2011 were extrapolated with the aid of simulated samples and using 2011 data rate-vs-luminosity curves, assuming as a rule-of-thumb that, due to the cross section increase, the rates for $\sqrt{s} = 8$ TeV and $L = 7 \times 10^{33}$ cm^{-2}s^{-1} would correspond to the rates for $\sqrt{s} = 7$ TeV and $L = 10^{34}$ cm^{-2}s^{-1}. It was estimated that the main di-photon trigger used during 2011 for both the $H \to \gamma\gamma$ search and SM cross section measurements, requiring two photon candidates with $E_{\mathrm{T}} > 20$ GeV passing loose identification requirements at EF and thus called in ATLAS's jargon `EF_2g20_loose`, would have produced an output rate of about 30 Hz, with a unique rate (*i.e.* the rate of events not collected by any other trigger, thus equal to the additional bandwidth required by this trigger) around 15 Hz. As the Higgs boson search was one of the priorities of the experiment, it was decided to keep the di-photon trigger unprescaled, but to increase the transverse momentum thresholds to reduce the trigger rate to a level that was judged to be acceptable while still keeping the trigger fully efficient for events passing the $H \to \gamma\gamma$ offline selection criteria, in which the two photons are required to have $E_{\mathrm{T}} > 40$ GeV and $E_{\mathrm{T}} > 30$ GeV, respectively. The main di-photon trigger for 2012 was thus chosen to be the `EF_g35_loose_g25_loose` one, with asymmetric E_{T} thresholds at 35 and 25 GeV.

Without additional di-photon triggers, this choice would have limited significantly the sample

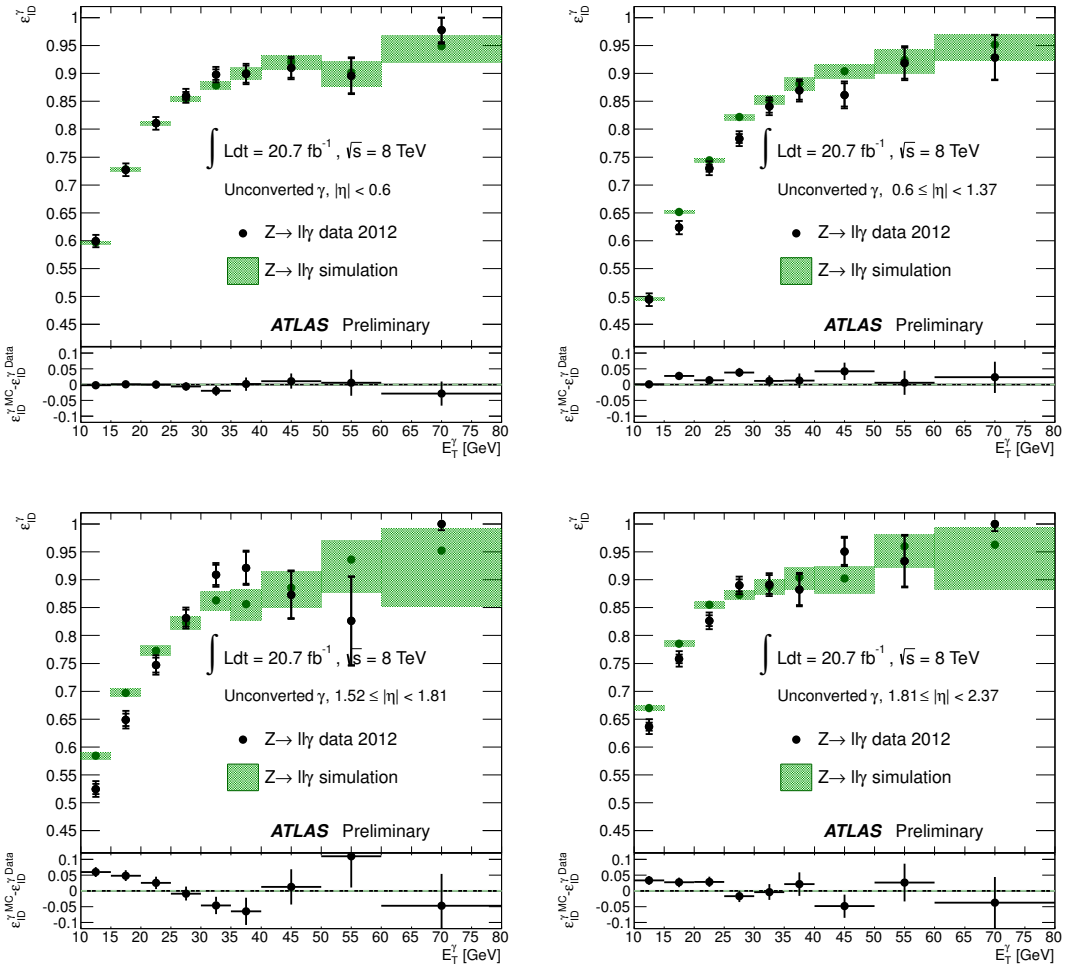

Figure 2.15: Unconverted photon identification efficiency as a function of the photon candidate E_{T}, as measured from $Z \to \ell\ell\gamma$ events. The data measurements are compared to the corresponding corrected simulation predictions. The inner and outer error bars represent the statistical and total uncertainties associated to the measurements, respectively. The green band represents the statistical uncertainty on the simulation. The bottom insets of each figure show the absolute difference between the data and the simulation.

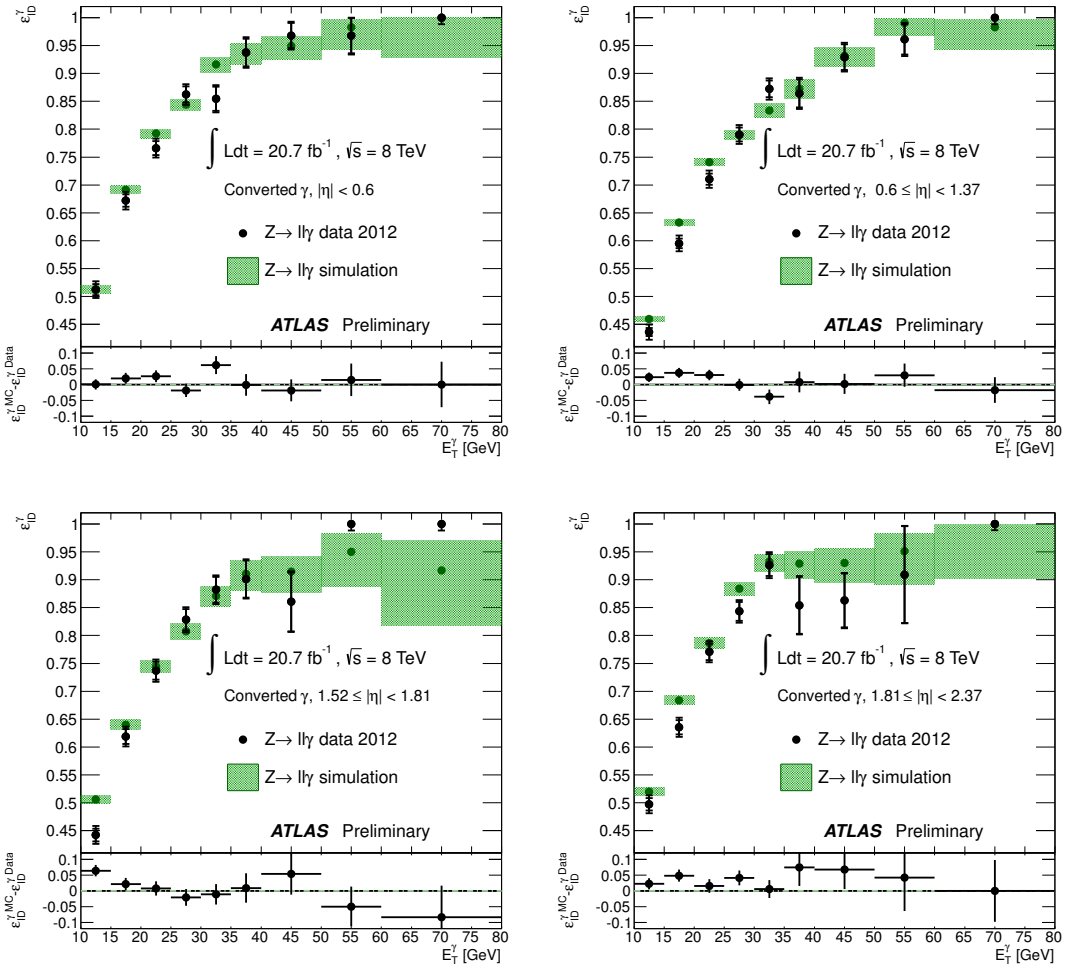

Figure 2.16: Converted photon identification efficiency as a function of the photon candidate E_{T}, as measured from $Z \to \ell\ell\gamma$ events. The data measurements are compared to the corresponding corrected simulation predictions. The inner and outer error bars represent the statistical and total uncertainties associated to the measurements, respectively. The green band represents the statistical uncertainty on the simulation. The bottom insets of each figure show the absolute difference between the data and the simulation.

of SM di-photon events (whose production cross section decreases with E_{T}) that could be collected and analyzed to measure their production cross section at $\sqrt{s} = 8$ TeV and compare it to the one measured at 7 TeV, where offline E_{T} thresholds of 25 and 22 GeV for the two photons had been used. A straightforward solution to this problem would have been to keep the `EF_2g20_loose` trigger in place (in addition to the `EF_g35_loose_g25_loose` one), but with a large (≈ 10) prescale factor to keep its unique rate below $2-3$ Hz. However, this would have reduced by a factor 10 the available di-photon statistics. Kun Liu, Yanwen Liu and myself proposed instead to define a "medium" photon trigger selection that would provide a higher jet rejection than the loose one by using relaxed requirements on one or more photon identification variables computed using the first layer of the ECAL, but that would still be fully efficient with respect to the offline tight identified photons. The proposal was accepted by the collaboration, and it was set as a goal to keep the unique rate of the new di-photon trigger below $2-3$ Hz, while trying at the same time to reduce the pile-up dependence of the (di-)photon trigger, in preparation for the large pile-up expected during the 2012 data-taking.

The study, which has been part of Kun Liu's qualification work to become an ATLAS author, has been carried on in December 2011 and the first 3 months of 2012, using data collected in 2011, as well as simulated samples of photons and QCD multi-jet events at $\sqrt{s} = 7$ and 8 TeV. Since the simulation, after correcting the photon shower shapes by small shifts as described previously, describes rather well the efficiencies measured in data, we rely on the corrected simulation for evaluating the impact of a trigger-level selection including also some layer-1 shower-shape variables. The pile-up dependence of the photon trigger is evaluated from the dependence of the trigger efficiency on the number of reconstructed primary vertices (N_{vtx}), using simulated photon-jet samples generated with an average number of interactions per bunch crossing equal to 20, 30 or 40. A large drop in efficiency (about 8%) with respect to reconstructed true photons is observed as N_{vtx} increases from 1 to 30, due to the use, at trigger-level, of two shower-shape variables, $R_{\mathrm{had}(,1)}$ and R_{η}, that are rather sensitive to pile-up. We have been able to reduce the pile-up dependency by almost 50% with a modest change (+8%, relative) in trigger rate by relaxing the trigger requirements on $R_{\mathrm{had}(,1)}$ and R_{η}, while tightening that on $w_{\eta 2}$.[5] In addition, we have investigated the change in trigger efficiency and rate by including layer-1 shower-shape variables in its selection. For this purpose we have decided not to use neither F_{side} nor $w_{s\,3}$, which are the variables that – exhibiting the smallest correlations with the isolation transverse energy of fake photons – are widely used in ATLAS to select, after reversing their requirements, control samples to study the isolation of background events. Among the other variables, we have chosen to use E_{ratio}, as its distribution for the signal is sharply peaked near one while for the background it has a long tail extending towards zero (see Figs. 1.12 and 1.13): a loose cut on E_{ratio} thus is highly efficient for the signal and provides additional rejection against hadrons. We have investigated several working points, corresponding to different requirements on E_{ratio} and thus different efficiencies for photons and rejections of fake candidates.

Figure 2.17 shows the efficiency for true photons and the inverse of the fake photon efficiency of a single-photon trigger with a 20 GeV threshold and various requirements on E_{ratio}: a "loose" cut at 0.3, a "medium" one at 0.6, a "tight" one around 0.75, and a "medium1" mixed scenario in which the threshold varies between 0.6 and 0.85 depending on the pseudorapidity interval. The brown dots correspond to the 2011 loose photon trigger while the black dots correspond to the new 2012 loose trigger where the pile-up dependence is reduced as described before. The results are obtained by selecting all the high-level trigger photon objects with $E_{\mathrm{T}} > 20$ GeV, $|\eta| < 2.47$ and matched to reconstructed photon candidates that are in turn matched or not to true prompt photons; the signal (background) efficiency, ε_s (ε_b), is computed as the fraction of all the trigger objects matched to a reconstructed true (fake) prompt photon candidate that pass the trigger-level requirements on the shower-shape variables. Based on these results, a decrease by about 1/3 for the single photon trigger rate, and thus by about 50% or better for the di-photon medium trigger, is expected using the "medium1" configuration. The same estimate of the rate reduction is confirmed using unbiased events collected in the full 2011 data with photon triggers in which the high-level trigger is in pass-through mode (and thus the shower-shape distributions are not biased, since no cut on them is applied at L1), as shown in Fig. 2.18. The trigger rate for a di-photon trigger

5. Eventually, cuts similar but not identical to our proposal were chosen in order to harmonize the photon trigger selection with the electron one. These requirements were used also for all the other photon "loose" triggers.

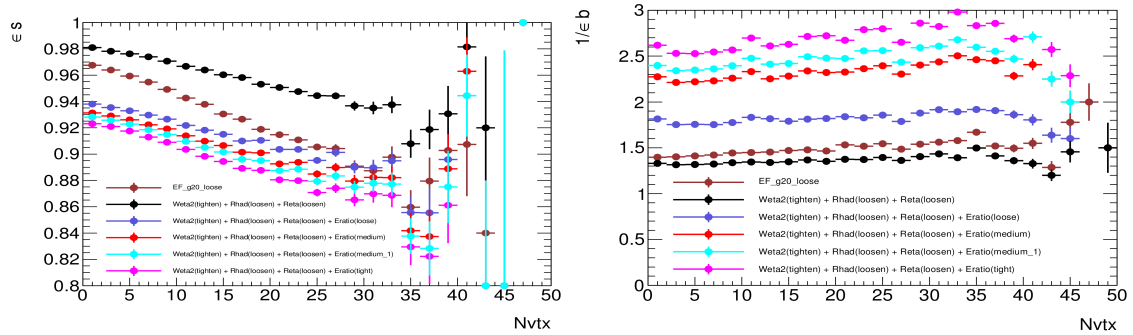

Figure 2.17: Efficiency of the the trigger-level shower-shape requirements, as a function of the number reconstructed primary vertices, for photon trigger objects matched to true (left) or fake (right) prompt photon candidates, in simulated photon-jet (left) or di-jet (right) events. In the right figure, the inverse of the efficiency is shown.

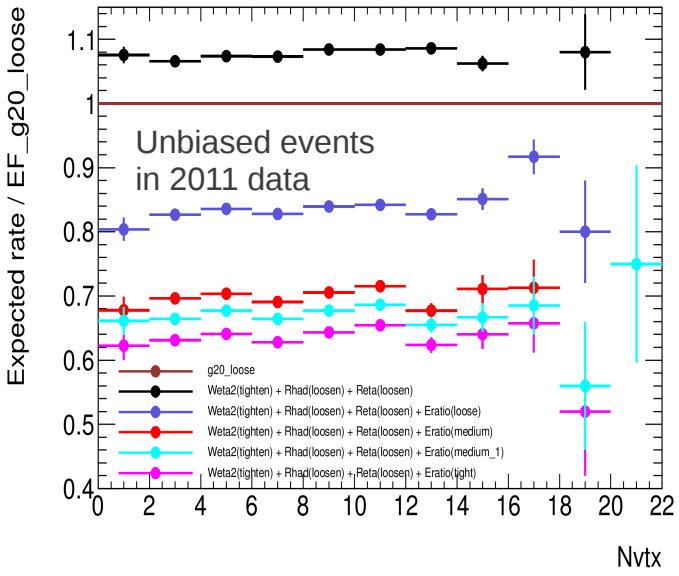

Figure 2.18: Ratio between the rate of various single-photon triggers (with E_{T} threshold equal to 20 GeV) proposed for the 2012 data-taking and the rate of the single-photon trigger with 20 GeV threshold used in 2011, based on data collected with the high-level photon trigger operating in pass-through mode.

(`EF_2g20_medium`) with symmetric E_{T} thresholds equal to 20 GeV and shower-shape requirements on the variables used by the loose trigger and the "medium1" selection on the E_{ratio} quantity is thus expected to be near 10 Hz for the 2012 data-taking, with a unique rate (computed checking in data and in simulated multi-jet events the fraction of events that pass `EF_2g20_medium` but not the other single-photon triggers nor the `EF_g35_loose_g25_loose` trigger) within 3 Hz. At the

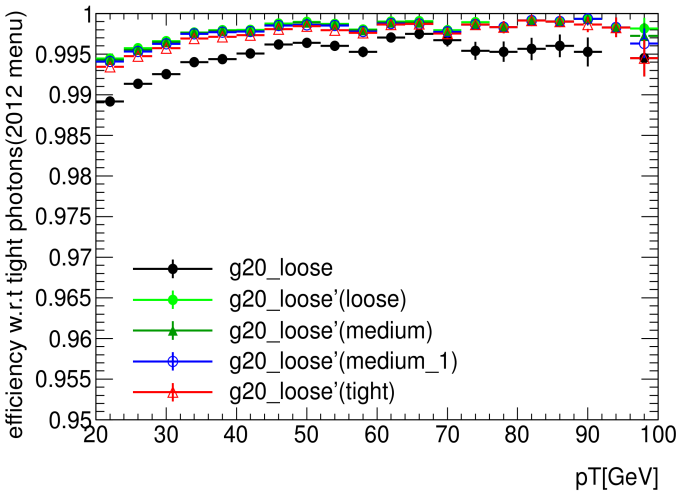

Figure 2.19: Efficiency of the trigger-level shower-shape requirements for photon objects matched to reconstructed true prompt photons passing the offline tight identification criteria in simulated photon-jet events, as a function of the photon E_{T}.

same time, the efficiency for photons passing the offline tight identification criteria is estimated with simulated photon-jet events to be larger than 99.3% for $E_{\mathrm{T}} > 20$ GeV for each of the four working points that have been investigated, as shown in Fig. 2.19. Based on these expectations, and on the consistency of the `EF_2g20_medium` trigger rate measured in the first 2012 data with the expectations, this trigger has been used to collect di-photon events down to photon $E_{\mathrm{T}} = 20$ GeV throughout the whole 2012.

2.4 Measurement of the photon trigger efficiency with 8 TeV data

The measurements of the SM prompt photon and $H \to \gamma\gamma$ cross sections use data collected with single- or di-photon triggers that apply looser photon selection criteria than those exploited in the offline identification algorithms and are thus highly efficient with respect to the final selection criteria. The single photon trigger efficiency is defined here as the probability that a tightly identified photon passes the trigger requirements; the di-photon trigger efficiency is defined similarly for photon pairs. The efficiency is computed for an inclusive sample of (di-)photons and also as a function of some reference quantities, like the photon pseudorapidity and transverse momentum. Evaluating this efficiency is part of the aforementioned cross section measurements, as the number of selected and identified signal events observed in the final data sample, N_{sel}, is related to the effective signal cross section (which may include the branching fractions of the decays of intermediate resonances) through the relation

$$N_{\mathrm{sel}} = \left(\int Ldt \right) \sigma_{\mathrm{eff}} \varepsilon_{\mathrm{rec}} \varepsilon_{\mathrm{ID}} \varepsilon_{\mathrm{trig}} \tag{2.13}$$

where $\int Ldt$ is the integrated luminosity of the analysed data, $\varepsilon_{\mathrm{rec}}$ is the signal acceptance and reconstruction efficiency in the detector (typically estimated from the simulation), $\varepsilon_{\mathrm{ID}}$ is the identification efficiency (relative to reconstructed isolated photons) computed with *in situ* techniques as described in Sec. 2.2, and $\varepsilon_{\mathrm{trig}}$ is the trigger efficiency, relative to reconstructed photons passing the tight identification.

Here I describe the *in situ* measurement of some single- and di-photon trigger efficiencies that Kun Liu, Yanwen Liu and I did in winter 2012 and spring 2013, using radiative decays of Z bosons in the 8 TeV data; more details are provided in Ref. [13]. I will not focus on other methods (like the "bootstrap" and the "tag-and-probe") that have also been used [13] and which I have studied in 2009, when contributing to their MC-based feasibility studies [123]. The idea is to use, for the measurement of the efficiency of photon triggers with rather low E_{T} thresholds (between 20 and 35 GeV), a clean sample of prompt, isolated photons of relatively low transverse momentum from $Z \to \ell\ell\gamma$ decays: to this purpose, the same selection as described in Sec. 2.2.5 is used. The trigger efficiency is measured with respect to photons passing the tight cut-based identification criteria and the calorimeter-isolation requirement $E_{\mathrm{T}}^{\mathrm{iso}} < 4$ GeV within a cone of radius 0.4 around the photon direction.

The measurement is based on the full ATLAS data set collected in 2012, retaining only those events in which all detectors are operational and have good data quality. To estimate systematic uncertainties, we also make use of 8 TeV Monte Carlo simulated samples of $Z \to \ell\ell$ (ten million each for $\ell = e, \mu$) and $Z \to \ell\ell\gamma$ (1.2 million each for the two lepton flavors) events, generated with SHERPA 1.4.1 [124, 125]. SHERPA, like PYTHIA, is a parton-shower Monte Carlo event generator, with its own implementation of the underlying event (AMISIC++), based on the multiple-parton interaction picture, and hadronization (AHADIC++), based on the cluster fragmentation model. Unlike PYTHIA it includes, in addition to the lowest-order matrix element for the various hard-scattering processes it implements, higher-order real-emission matrix elements; for the studies presented here, up to five partons are generated together with the Z in the $Z \to \ell\ell$ samples and up to three partons are generated with the Z in the $Z \to \ell\ell\gamma$ ones.

Photons are required to have transverse energy $E_{\mathrm{T}} > 10$ GeV and pseudorapidity $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$. ISR events, which are largely affected by the Z+jets background, are suppressed by requiring $70 < m_{\ell\ell\gamma} < 100$ GeV and $40 < m_{\ell\ell} < 83$ GeV. The two-dimensional (E_{T}, η) distribution of the photon candidates in $Z \to \ell\ell\gamma$ events selected in data is shown in Fig. 2.20. Their one-dimensional transverse momentum and pseudorapidity distributions of photon candidates are shown in Fig. 2.21. After all requirements, 79452 tight photon candidates with $E_T > 10$ GeV are selected in the full 2012 data set.

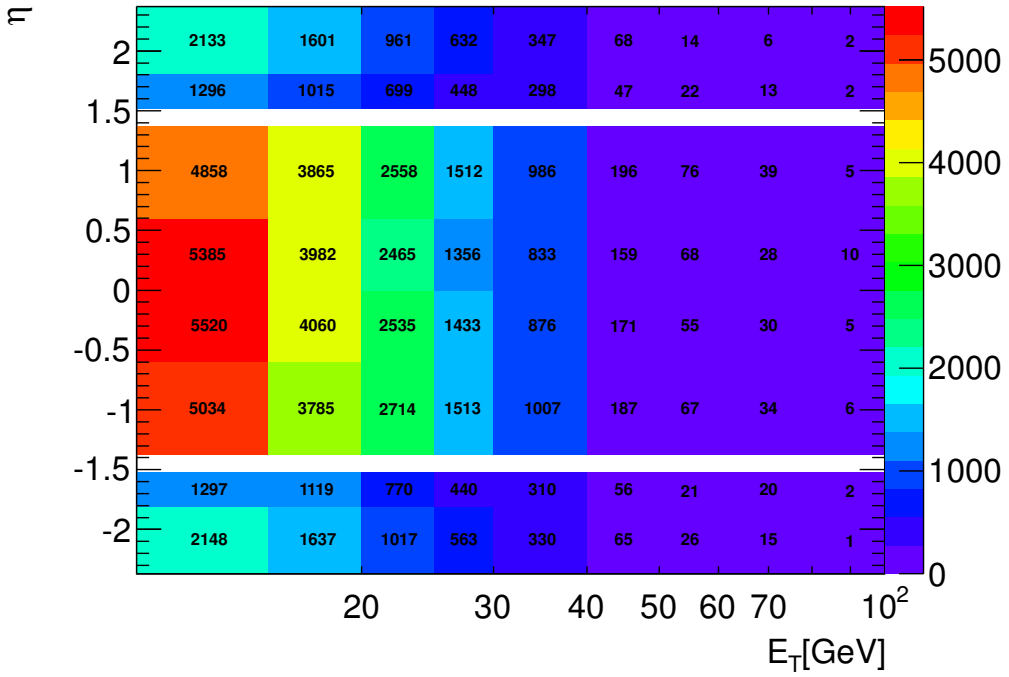

Figure 2.20: 2-D distribution of E_{T} and η for photon candidates selected in data after the full $Z \to \ell\ell\gamma$ selection is applied.

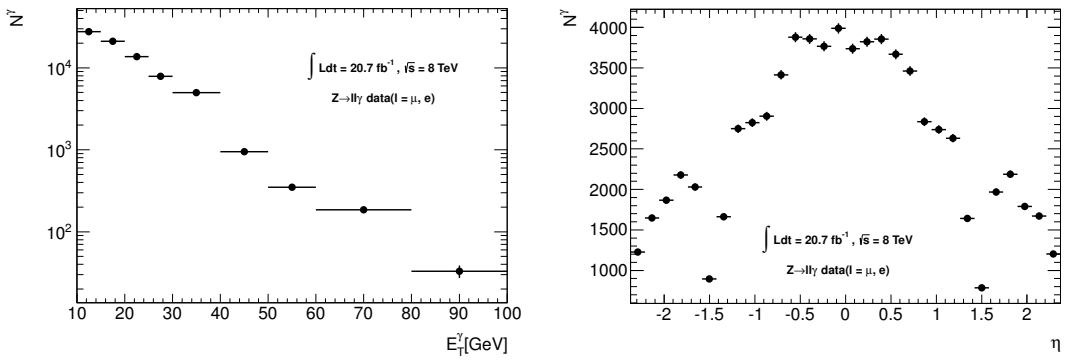

Figure 2.21: Distributions of E_{T} (left) and η (right) for photon candidates selected in data after the full $Z \to \ell\ell\gamma$ selection is applied.

The photon purity of the selected sample, estimated through the signal+background fit to the three-body invariant mass described in Sec. 2.2.2 using the aforementioned simulated samples to determine the signal and background distributions, is around 96% for photon transverse momenta between 10 and 15 GeV, and increases to 99% and above for $E_{\mathrm{T}} > 15$ GeV. The trigger efficiency curves, as a function of the photon pseudorapidity and transverse momentum, are determined from the fraction of all the offline photon candidates passing the tight identification and the isolation requirement that are within a distance $\Delta R = 0.15$ in $\eta - \phi$ from a photon trigger object. The efficiency as a function of the photon transverse momentum for the `EF_g20_loose` and the

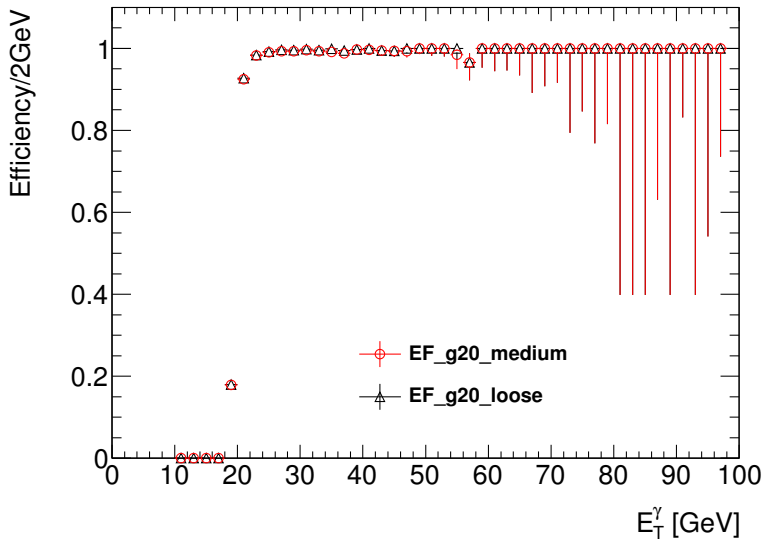

Figure 2.22: Efficiency as a function of the photon transverse momentum for the `EF_g20_loose` and the `EF_g20_medium triggers`. The efficiency is measured with respect to offline photon candidates with pseudorapidity $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$, passing the tight identification and the isolation requirement $E_{\mathrm{T}}^{\mathrm{iso}} < 4$ GeV.

`EF_g20_medium` triggers, *i.e.* single-photon triggers requiring photon candidates with $E_{\mathrm{T}} > 20$ GeV and passing loose or medium1 identification criteria, is shown in Fig. 2.22 for photons with pseudorapidity $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$. The efficiency of the same triggers as a function of the photon pseudorapidity, for transverse momenta $E_{\mathrm{T}} > 25$ GeV, is shown in Fig. 2.23. The integrated efficiency for photons with $E_{\mathrm{T}} > 25$ GeV and $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$, passing the tight identification criteria and the isolation requirement $E_{\mathrm{T}}^{\mathrm{iso}} < 4$ GeV, is

- `EF_g20_loose`: $(99.62^{+0.05}_{-0.06})\%$

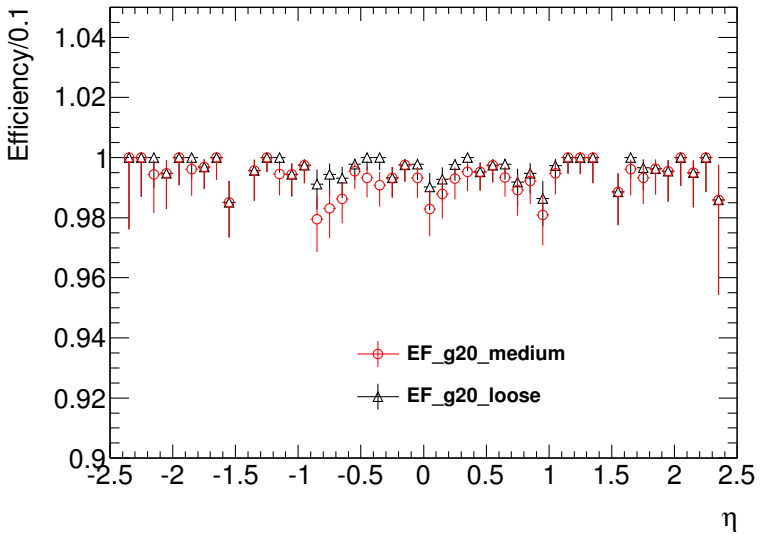

Figure 2.23: Efficiency as a function of the photon pseudorapidity for the `EF_g20_loose` and the `EF_g20_medium` triggers. The efficiency is measured with respect to offline photon candidates with transverse momentum $E_{\mathrm{T}} > 25$ GeV, passing the tight identification and the isolation requirement $E_{\mathrm{T}}^{\mathrm{iso}} < 4$ GeV.

- `EF_g20_medium`: $(99.33^{+0.07}_{-0.08})\%$

where the uncertainty is statistical, only.

Systematic uncertainties originate from the small background contamination of the selected control sample, if the trigger efficiency is different between photons and fake candidates. However, since the photon purity is more than 99% above 15–20 GeV and the fake candidates passing the tight identification criteria are rather similar to prompt photons in the calorimeter, the bias is expected to be very small. A systematic uncertainty due to the presence of a small fake photon contamination is estimated by repeating the measurement after selecting $Z \to \ell\ell\gamma$ candidates with different requirements on the three-body mass $m_{\ell\ell\gamma}$: both the lower and upper limit are varied independently by ± 10 GeV, and the photon purity of the selected candidates varies by as much as 1.3%, thus covering the estimated difference between the purity of the default control sample and a 100% pure photon sample. The trigger efficiency determined from the alternative control samples varies by at most $\pm 0.05\%$ for both the `EF_g20_loose` and the `EF_g20_medium` trigger. The presence of a QCD background in addition to the Z+jets background that was considered so far in the $m_{\ell\ell\gamma}$ fit to estimate the photon purity is assessed by using $\ell\ell\gamma$ events in which the leptons have the same charge to determine the shape of the $m_{\ell\ell\gamma}$ distribution of this background, and using the sideband regions $40 < m_{\ell\ell\gamma} < 65$ GeV and $105 < m_{\ell\ell\gamma} < 120$ GeV in events in which the leptons have opposite charge to estimate the overall normalisation. The QCD background yield in the selected control sample is estimated to be small (around 20 events), with an impact on the efficiency evaluated to less than 0.01%. Finally, a possible systematic uncertainty associated to the matching criterion between the offline photon candidate and the photon trigger objects is estimated by varying the ΔR matching distance between 0.1 and 0.2. The impact on the measured efficiency is negligible.

Including the previous uncertainties, the integrated efficiency for photons with $E_{\mathrm{T}} > 25$ GeV and $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$, passing the tight identification criteria and the isolation requirement $E_{\mathrm{T}}^{\mathrm{iso}} < 4$ GeV, is

- `EF_g20_loose`: $(99.62^{+0.05}_{-0.06}\ (\mathrm{stat})^{+0.05}_{-0.05}\ (\mathrm{syst}))\%$
- `EF_g20_medium`: $(99.33^{+0.07}_{-0.08}\ (\mathrm{stat})^{+0.05}_{-0.05}\ (\mathrm{syst}))\%$

The efficiency of the di-photon triggers is estimated by multiplying the efficiencies of two single-

photon trigger efficiencies: the correlations between the two efficiencies can be neglected for di-photon final states where the photon calorimeter clusters (and trigger objects) are sufficiently separated so as not to overlap, which is guaranteed by the photon isolation requirement and is typically enforced offline by a $\Delta R_{\gamma\gamma} > 0.4$ requirement. With this method, the di-photon trigger efficiency is estimated to be:

- `EF_2g20_medium`: $(98.66^{+0.14}_{-0.16}\ (\text{stat})^{+0.10}_{-0.10}\ (\text{syst}))\%$
- `EF_g35_loose_g25_loose`: $(99.48^{+0.19}_{-0.29}\ (\text{stat})^{+0.10}_{-0.10}\ (\text{syst}))\%$
- `EF_g30_medium_g20_medium`: $(98.71^{+0.22}_{-0.26}\ (\text{stat})^{+0.10}_{-0.10}\ (\text{syst}))\%$

for events in which the two photons satisfy the following criteria:

- transverse momenta greater than 25 GeV in the first case, 40 and 30 GeV in the second one, 35 and 25 GeV in the third one,
- $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$,
- tight identification,
- isolation $E_{\text{T}}^{\text{iso}} < 4$ GeV.

Finally, as these efficiencies are estimated using photon control samples from $Z \to \ell\ell\gamma$ decays, which have definite E_{T} and η distributions, an additional systematic uncertainty is estimated by comparing the di-photon trigger efficiency in simulated $H \to \gamma\gamma$ decays (for a Higgs boson mass of 125 GeV, applying the offline $H \to \gamma\gamma$ selection criteria) and the one estimated on MC from simulated $Z \to \ell\ell\gamma$ with the method previously described. Although the photon kinematic in these two decays is rather different, the two efficiencies differ by less than 0.1%, and a 0.1% systematic uncertainty on the di-photon trigger efficiency due to the di-photon kinematic distribution is thus assigned.

Chapter 3

Measurement of the production cross sections of isolated prompt photons

Standard Model processes producing prompt photons (photon-jet and di-photon events) provide clean samples to test perturbative QCD predictions and to extract information about some parton distribution functions of the proton. In addition, these processes are the main backgrounds for the search of Higgs boson decays to di-photon final states, and measuring their cross sections supplies useful information to tune the programs used in the simulation to generate background samples for the optimisation of the $H \to \gamma\gamma$ selection criteria.

The measurement of the production cross sections of isolated prompt photons, either inclusively, in association with jets, or in pairs, and their comparison to the Standard Model theoretical predictions, has been one of the main data analysis activities I have been involved with in the past years. These activities began in spring 2010, soon after finishing the photon identification optimization work and before the ICHEP conference, when with Guillaume Unal I coordinated and led the analysis of the first 15 nb^{-1} of 7 TeV ATLAS data, in order to demonstrate the ability of the ATLAS detector to reconstruct prompt photons and to distinguish them from background events, as documented in Ref. [126]. This gave me the chance to work with a group of ATLAS physicists with a huge experience of the ECAL and of prompt photon physics, as well as several young and bright Ph.D. or post-doc students and their supervisors.

After ICHEP 2010, with more data available, I turned to the first measurement of the isolated prompt photon production cross section, using 0.9 pb^{-1}. Again, I had the opportunity to lead this effort, together with Mike Hance. The measurement, completed in winter 2010 and published at the beginning of 2011 [14], was based on the same background-subtraction technique developed in [126], and contained in addition a complete evalutation of the reconstruction, identification and trigger efficiencies, needed for the cross section measurement.

In the following months of 2011, while focusing mostly on the photon-jet cross section determination [16], which I performed with a few colleagues in Milano (in particular, Leonardo Carminati and Iro Koletsou) and was eventually published in May 2012, I also contributed to the measurements, using the full 2010 ATLAS dataset, of the production cross sections of prompt photons [15] and of photon pairs [17], coordinated respectively by Thomas Koffas, Martin Tripiana and Mark Stockton, and by Marcello Fanti and Sandrine Laplace.

Between October 2011 and October 2012 I was convener of the ATLAS SM Direct Photon (SMDP) working group, a small analysis team of a few ($\lesssim 20$) but dedicated people working on Standard Model photon production cross sections. During that year I supervised the measurement of the prompt photon cross section with a much larger dataset (4.7 fb^{-1} of data collected in 2011), described in [127], and other studies whose results are not yet public, though my main efforts were the conclusion of the photon-jet cross section measurement and, with some colleagues of the LPNHE, LAPP and Milano groups (in particular, Lydia Roos, Mayuko Kataoka and Remi Lafaye), the measurement of the di-photon production cross-section with the full 2011 dataset [18],

published in January 2013.

In this chapter I will summarize the main measurements that are published and to which I contributed the most, their results, and the consequences on the gluon and light quark PDFs.

3.1 Theoretical overview

In hadronic collisions, prompt photons are photons in the final state that do not originate from hadron decays. Due to the pointlike nature of the quark-photon vertex, measuring the production cross section of prompt photons in hadron-hadron interactions with large momentum transfer has been proposed since more than thirty years as a clean source of information on the hard-scattering dynamics of quarks and gluons [128, 129, 130]. In comparison, the theoretical interpretation of inclusive hadron production cross sections is complicated by hadron form-factors and fragmentation functions, while jet production cross sections, which are more directly related to the underlying parton-level QCD cross sections than those of single hadrons and are enhanced – with respect to prompt photon production – by a factor α_s/α, are affected by ambiguities related to the experimental jet reconstruction algorithm, important energy scale, energy resolution and direction uncertainties, and a significantly larger number of Feynman diagrams that contribute to the final state and need to be included in the theoretical calculation. Moreover, as the photon couples to electric charge while quarks and gluons interact via color charges, prompt photon production gives complementary information to that provided by jet production.

A detailed theoretical overview of prompt photon production is given in Refs. [129, 130, 131], which are briefly summarized here. The simplest Feynman diagrams describing the Born-level production of prompt photons are the QCD Compton scattering, $qg \to q\gamma$ (Fig. 3.1a), and quark anti-quark annihilation, $q\bar{q} \to g\gamma$ (Fig. 3.1b). Both diagrams are of order $\mathcal{O}(\alpha\alpha_s)$. In the case of large momentum transfers compared to the QCD scale $\Lambda_{\mathrm{QCD}} \approx 0.2$ GeV, and thus for events where the photons are produced with transverse energy $E_{\mathrm{T}} \gg \Lambda_{\mathrm{QCD}}$, the partons are in the "asymptotic freedom" regime in which the running coupling α_s becomes small enough to justify the use of perturbative techniques. To make a precise test of pQCD, however, it is necessary to go beyond the leading-order parton-level cross section and to take into account the following issues:

- The incoming beams in the experiment are composed of hadrons, not of partons. Starting from the elementary parton-level cross sections, the hadron-hadron cross section is thus computed, with the help of the parton model: the elementary cross sections are calculated for incoming partons carrying fractions x_a and x_b of the colliding hadron momenta and convoluted with the parton distribution functions $F_a(x_a)$ and $F_b(x_b)$, describing the probability for a parton of type a (b) to have a fractional momentum x_a (x_b). An integration over all possible values of x_a and x_b and a sum over all parton types a and b ($a, b = q, \bar{q}, g$) are performed.
- QCD corrections from higher-order diagrams have to be calculated. At next-to-leading order, $\mathcal{O}(\alpha\alpha_s^2)$, they include the subprocesses $qg \to qg\gamma$ (Fig. 3.1c), $q\bar{q} \to gg\gamma$ and $q\bar{q} \to q\bar{q}\gamma$ (Fig. 3.1d), as well as the virtual corrections to the Born-level processes (as for instance those shown in Fig. 3.1e). As soon as the contributions from higher-order diagrams are considered, divergences start to plague the theoretical calculation. A class of logarithmic divergences, referred to as "mass singularities", happens as a consequence of the masslessness of the partons when including for instance the contribution from diagrams in which an additional gluon, radiated from a quark, is present in the final state of the Compton process and is collinear with the emitting quark. These singularities, however, according to the factorization theorem, are universal, in the sense that they affect in the same way all subprocesses containing a given species of parton, and can be separated out and absorbed into the parton distribution functions, by fixing a momentum scale ("factorization" scale) μ_F below which the parton scattering process is considered a radiative correction to the incoming partons and is accounted for by replacing the "bare" structure function $F_a(x_a)$ with the "dressed" function $F_a(x_a, \mu_F^2)$.
  Similar divergences arise when the particle emitted collinearly from a hard parton participating to the short-distance subprocess is the photon: this is the so-called "fragmentation"

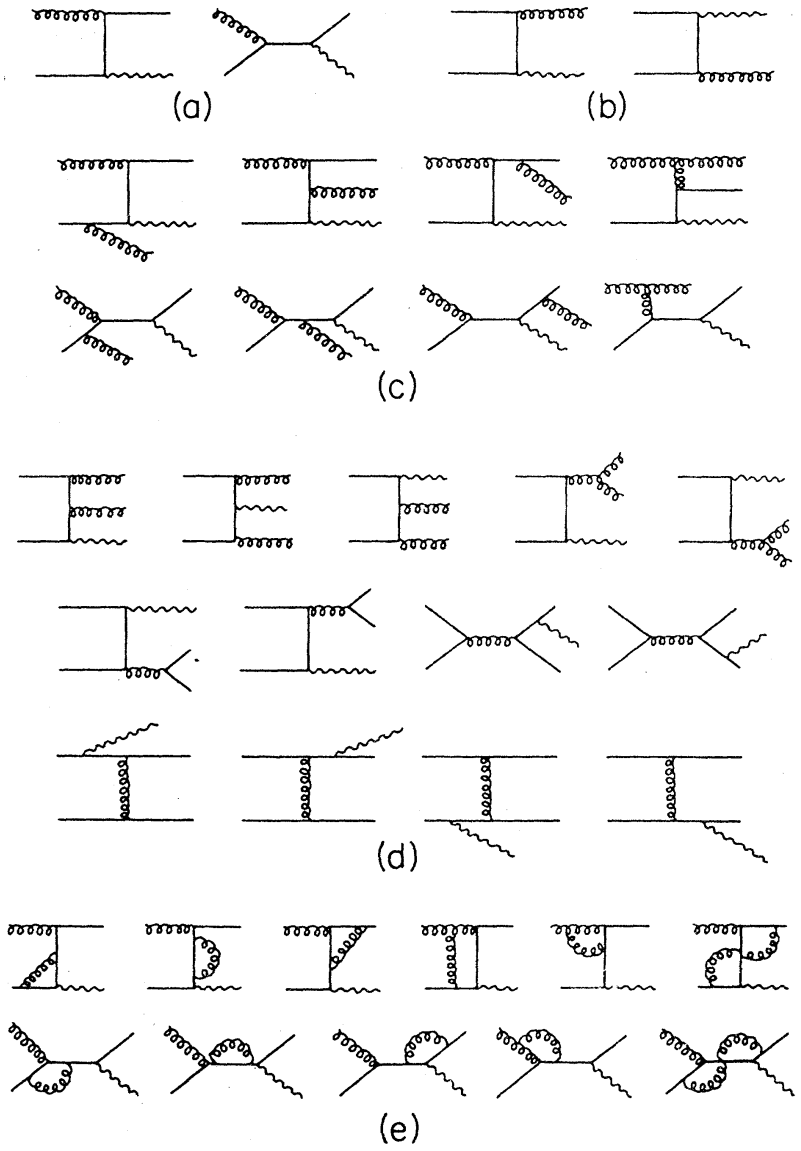

Figure 3.1: Selection of Feynman diagrams for prompt photon production: (a) $\mathcal{O}(\alpha\alpha_s)$ Compton subprocess, $qg \to q\gamma$ (and $\bar{q}g \to \bar{q}\gamma$); (b) $\mathcal{O}(\alpha\alpha_s)$ annihilation subprocess, $q\bar{q} \to g\gamma$; (c) $\mathcal{O}(\alpha\alpha_s^2)$ subprocess $qg \to \gamma qg$ (and $\bar{q}g \to \gamma\bar{q}g$); (d) $\mathcal{O}(\alpha\alpha_s^2)$ subprocesses $q\bar{q} \to \gamma gg$ and $q\bar{q} \to \gamma q\bar{q}$; (e) one-loop graphs for the Compton subprocess that contribute to the $\mathcal{O}(\alpha\alpha_s^2)$ calculation [129].

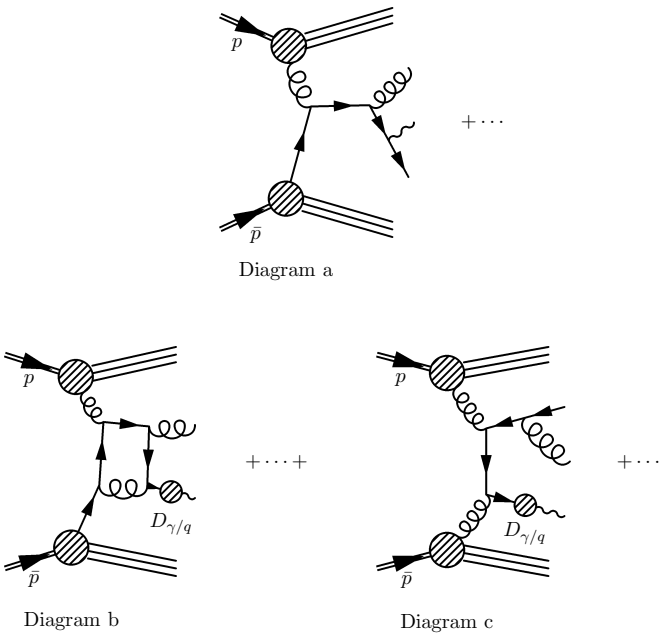

Figure 3.2: Examples of diagrams contributing at LO (top) and NLO (bottom) to the fragmentation prompt-photon production process [131].

or bremsstrahlung contribution. An example of leading-order diagram of the fragmentation type is shown in Fig. 3.2a. As in the previous case, these singularities are universal and can be factorized at all orders in α_s and absorbed into quark and gluon fragmentation functions of the photon, $D^{\gamma}_{q/g}(z, \mu_f^2)$, defined at a "fragmentation" scale μ_f of the order of the hard scale of the process. z is the fraction of parton momentum carried by the photon. For μ_f large with respect to ≈ 1 GeV, these functions are roughly of order $\alpha/\alpha_s(\mu_f^2)$ and the fragmentation contribution is thus of the same order in α_s as the "direct" mechanism where the photon participates to the hard subprocess; a NLO calculation of the prompt photon production cross section should thus include the NLO corrections to the fragmentation contribution, exemplified in Fig. 3.2b,c. [1]

Additional, "ultraviolet" singularities arise from diagrams with internal gluon lines running in loops. These can be regulated using a renormalization procedure based for instance on dimensional regularization, and absorbed into a "running" coupling constant $\alpha_s(\mu_R^2)$ which depends on a "renormalization scale" μ_R.

The prompt photon production cross section is thus [131]:

$$d\sigma \equiv d\sigma_{\rm dir} + d\sigma_{\rm frag} = \sum_{a,b=q,\bar{q},g} \int dx_a dx_b \; F_a(x_a, \mu_F^2) F_b(x_b, \mu_F^2) \times$$
$$\left[d\hat{\sigma}^{\gamma}_{ab}(x_a, x_b, \mu_R, \mu_F, \mu_f) + \sum_{c=q,\bar{q},g} \int_{z_{\rm min}}^{1} \frac{dz}{z^2} d\hat{\sigma}^{c}_{ab}(x_a, x_b, z, \mu_R, \mu_F, \mu_f) D^{\gamma}_{c}(z, \mu_f^2) \right] , \quad (3.1)$$

where $d\hat{\sigma}^{\gamma}_{ab}$ describes the direct production of a photon in the elementary process $a + b \to \gamma + d$, while $d\hat{\sigma}^{c}_{ab}$ describes the production of a parton c in a hard collision, $a + b \to c + d$; both cross sections are evaluated for $\alpha_s = \alpha_s(\mu_R^2)$. The center-of-mass energy squared of the incoming partons is $\hat{s} = x_a x_b s$.

As can be seen in Eq. 3.1, the cross section depends on the distribution functions of the interacting partons in the proton. In particular, at the LHC, the dominating processes are, for

1. It should be noted that the distinction between the fragmentation and the direct process is unphysical beyond LO, as the same diagram can contribute to either one, depending on the cut-off set by the fragmentation scale. The only physical quantity is the total cross section resulting from the sum of the direct and fragmentation contributions.

direct production, the QCD Compton, $qg \to q\gamma$, and, among the fragmentation ones, the $gg \to q\bar{q}$ and $gq \to gq$ scattering. For this reason, the prompt photon production cross section can be particularly interesting for PDF fits as it provides direct information on the proton gluon PDF $g(x, Q^2)$ [66, 67] at low x, which is otherwise only indirectly constrained through scaling-violation effects in deep-inelastic-scattering ep collisions ($\partial F_2^p(x, Q^2)/\partial \log Q^2$) [132] and from the momentum-sum rule. Since a significant fraction of the hard-scattering events at the LHC is due to gluon interactions and that gluon-gluon fusion is the dominant production process for SM Higgs bosons but also for other signatures like $t\bar{t}$ and di-jet events, improving the knowledge of $g(x, Q^2)$ is extremely important. Given the LHC center-of-mass energy and the large energy and rapidity

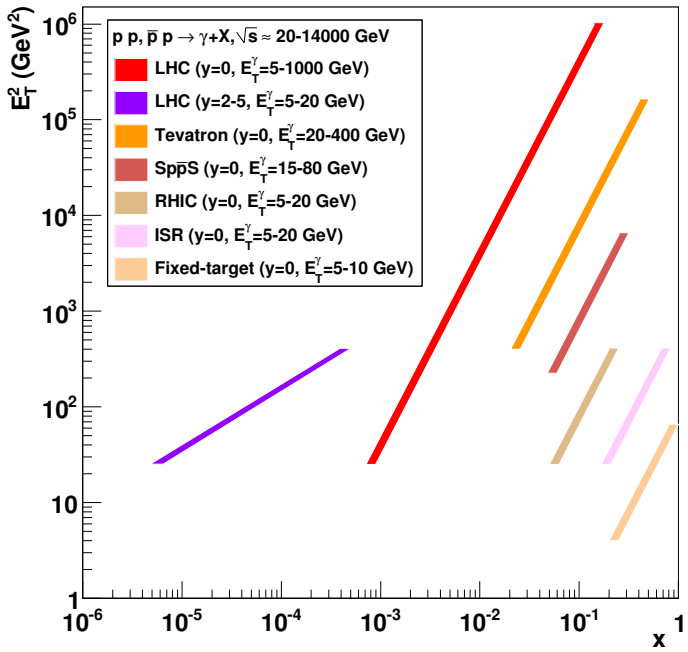

Figure 3.3: Kinematical region probed by pre-LHC prompt photon measurements at fixed-target (Fermilab) and collider (ISR, RHIC, SppS, Tevatron) energies, and expected range probed at the LHC ($\sqrt{s} = 14$ TeV) at central ($y = 0$) and forward ($y = 2 - 5$) photon rapidities [133].

acceptance of the LHC detectors, the kinematic region explored by prompt photon measurements can cover a wide range of the parton momentum fraction, given in the leading-order approximation by $x_\pm = \frac{E_T^\gamma}{\sqrt{s}}(e^{\pm\eta^\gamma} + e^{\pm y^{\text{jet}}})$, and of the hard-scattering scale, $Q^2 = (E_T^\gamma)^2$, as shown in Fig. 3.3. The prompt photon cross section is also sensitive to some extent to the quark PDFs, in particular the u one: since the photon couples to electric charge, the ratio $u : d$ for the Compton process in pp collisions is expected to be around 8, with a factor of 4 coming from the square of the quark charge and a factor of 2 from the fact that there are two valence u quarks and one valence d quark in the proton.

The calculation of the prompt di-photon production cross section is carried on in a similar way as that of the single photon production; a detailed overview is given in Ref. [134]. One starts with the LO calculation of the Born parton-level cross section, given by the $\mathcal{O}(\alpha^2)$ $q\bar{q} \to \gamma\gamma$ "direct" process (Fig. 3.4a), and then proceeds to compute the QCD NLO corrections $\mathcal{O}(\alpha_s\alpha^2)$ from the $q\bar{q} \to \gamma\gamma g$ and $gq \to \gamma\gamma q$ processes and corresponding virtual diagrams (e.g. Fig. 3.4b,c). This introduces a collinear singularity which is factorized at all orders in α_s and absorbed into the quark and gluon fragmentation functions to a photon. The collinear divergent part of the previous diagrams thus yields the leading-order contribution of the "single fragmentation" process (also called "bremsstrahlung" contribution) of Fig. 3.4d, which is of the same order $\mathcal{O}(\alpha^2)$ as the Born diagram since the fragmentation functions behave as α/α_s. For a full NLO treatment the NLO corrections to the single fragmentation process (Figs. 3.4e,f) should thus be calculated. This yields in turn the leading-order contribution of the "double fragmentation" process (Fig. 3.4g), in which

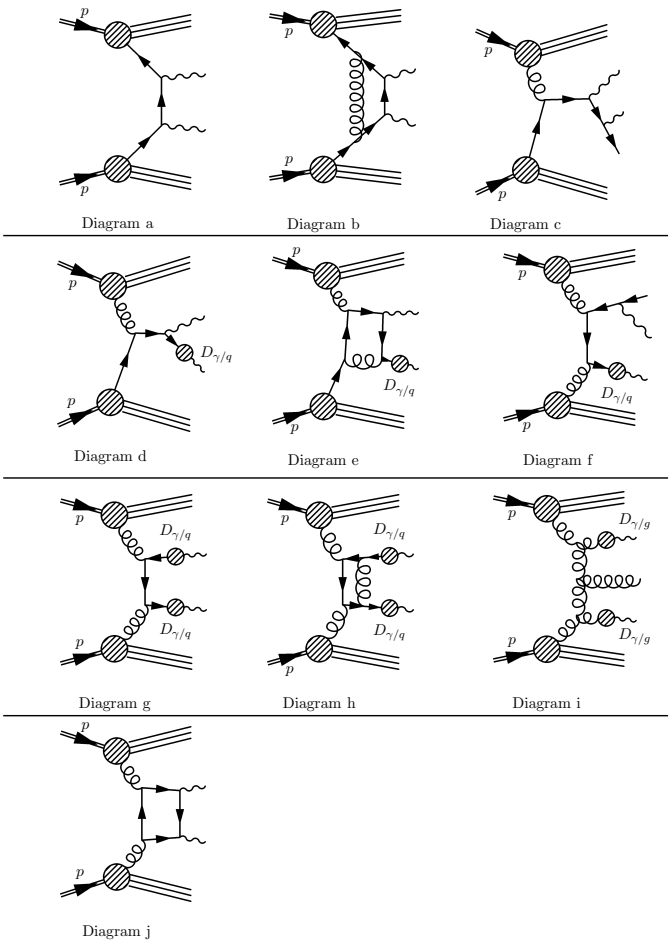

Figure 3.4: Examples of diagrams contributing to the di-photon production cross-section: direct (first row), single fragmentation (second row), double fragmentation (third row), and box (fourth row) diagrams. [134]. Each of the first three rows shows examples of the LO diagram (left) and of the NLO corrections (center, right).

both photons result from the collinear fragmentation of a hard parton. Finally, NLO corrections to the double fragmentation contribution (e.g. Fig. 3.4h,i) have to be computed. Beyond NLO, additional diagrams need to be evaluated, like the "box" gluon fusion diagram of Fig. 3.4j and the NNLO corrections to the previous diagrams, including the collinear finite part of the $2 \to 4$ processes $gg \to \gamma\gamma q\bar{q}$, $q\bar{q} \to \gamma\gamma gg$, $q\bar{q} \to \gamma\gamma q\bar{q}$, $gq \to \gamma\gamma gq$. The parton-level cross sections are computed as a function of the incoming parton momentum fractions x_a and x_b, weighted by $F_a(x_a)F_b(x_b)$ and an integral over x_a and x_b is performed.

3.2 Theoretical calculations

3.2.1 Inclusive photon and photon-jet production cross section

The expected production cross section of an isolated prompt photon in proton collisions at $\sqrt{s} = 7$ and 8 TeV is computed with the Jetphox Monte Carlo program [131, 135, 136], a parton-level photon-jet(s) event generator which implements a full NLO QCD calculation of both the direct and the fragmentation contributions. In addition, given the large gluon luminosity at the LHC [137], the authors of Jetphox have calculated the NNLO, $\mathcal{O}(\alpha\alpha_s^3)$, gluon-gluon box diagram $gg \to \gamma g$, which however they found to give a negligible contribution to the total cross section. We use the NLO BFG-II photon fragmentation function [138], which is in agreement with ALEPH and OPAL data, and various PDF sets, provided by the `LHAPDF` package [139]. The nominal renormalization (μ_R), factorization (μ_F) and fragmentation (μ_f) scales are set to the photon transverse energy E_{T}^{γ}. Jets of partons are reconstructed by using an anti-k_T algorithm with a radius parameter $R = 0.4$. The total transverse energy carried by the partons inside a cone of radius $R = 0.4$ in the $\eta - \phi$ plane around the photon direction is required to be less than a certain threshold, typically matched to the experimental requirement as described later. This requirement significantly suppresses the contribution to the total cross section from the fragmentation process, from $50 - 60\%$ down to typically less than $10 - 15\%$, as illustrated in Fig. 3.14. The measurement of the inclusive isolated photon production cross section, $pp \to \gamma + X$, described in Sec. 3.3, is compared to the calculation obtained with Jetphox 1.2.2 using the CTEQ 6.6 [140] PDFs without applying any selection on the jet momenta and directions, while the measurement of the production of isolated photons in association with jets, $pp \to \gamma + jet + X$, described in Sec. 3.5, is compared to the calculation obtained with Jetphox 1.3, using the CT10 NLO [141] PDFs, after applying to the parton jets the same transverse momentum and rapidity criteria applied in the measurement to the reconstructed hadron jets.

Jetphox calculates the NLO cross section for partons in the final state, thus it does not include the effects of hadronization nor of the underlying event and of pileup. While the ambient-energy density corrections to the photon isolation (Sec. 1.6.5) are expected to remove most of these effects from the photon isolation transverse energy, residual differences between the photon particle-level and parton-level isolations and between particle-level and parton-level jets can still be present. Simulated photon-jet events produced with various parton-shower Monte Carlo generators are thus used to compute the bin-by-bin ratios of the differential generator-level cross sections with and without hadronization and underlying event, and the predicted differential parton-level cross sections are multiplied by these ratios. These corrections usually deviate from unity by at most a few %.

The systematic uncertainties on the QCD cross sections computed with Jetphox are estimated in the following way:

- The uncertainty related with the arbitrary choice of the energy scales is evaluated by fixing any two scales to the nominal value and varying the third between 0.5 and 2.0 times the nominal value. In addition the effect of the coherent scale variations where all three scales are varied together is also taken into account. The envelope of the values obtained with the different scale configurations is taken as a systematic uncertainty.
- The uncertainty due to the limited PDF accuracy is obtained by varying each of the eigenvalues of the PDFs by $\pm 1\sigma$ and summing in quadrature separately positive and negative variations of the cross section. As a cross-check, the theoretical cross section is also recomputed using alternative PDF sets (MSTW, NNPDF); the typical change is of the order of a

few % and is within the uncertainty band from the PDF eigenvalue variations.

- The uncertainty due to the hadronization and underlying-event corrections is estimated as the envelope of the correction factors obtained from alternative hadronization and underlying-event models. It is of the order of 1-2%.

The results of the photon-jet cross section measurement are also compared to the generated particle-level spectra predicted by two leading-order parton-shower MC generators used in the ATLAS full simulation, PYTHIA and SHERPA. PYTHIA implements the LO matrix elements for prompt photon production in $qg \to q\gamma$ and $q\bar{q}$ to $q\gamma$ ($gg \to g\gamma$ is also included but gives a negligible contribution). In addition, "fragmentation-like" events in which the photon is radiated off a parton are included by enhabling QED radiation off quarks and generating all tree-level $2 \to 2$ QCD processes; only events where a photon is radiated are then retained and passed to the full detector simulation. Unlike JETPHOX, PYTHIA includes a model for the underlying event (based on the multiple-parton interaction model) and for the hadronization of partons in the final state (based on the Lund string model). SHERPA has features similar to those of PYTHIA, with different underlying-event and hadronization models and the additional inclusion of higher-order real-emission matrix elements; for the studies presented here, up to three partons in addition to the leading-p_{T} one and the photon are considered, and the photon is required to have an angular separation $\Delta R > 0.3$ with respect to the partons from the matrix element. Both PYTHIA and SHERPA are expected to underestimate the total cross section, because of the missing NLO (and higher-order) contributions. In the parton-shower simulated samples, truth-particle jets (or truth jets) are reconstructed using the anti-k_t algorithm with a radius parameter $R = 0.4$ on all the particles with proper lifetime longer than 10 ps, including photons but excluding muons and neutrinos, and the leading-p_{T} truth jet is selected among those with axis separated from the photon direction by $\Delta R > 0.3$. The photon isolation transverse energy is computed at truth-particle-level from the true four-momenta of the generated particles (excluding muons and neutrinos) inside a cone of radius 0.4 around the photon direction and is corrected for underlying event by subtracting the average energy density multiplied by the area of the isolation cone, where the energy density is estimated from the soft truth-particle jets in the event similarly to what is done experimentally (Sec. 1.6.5).

3.2.2 Di-photon production cross section

The di-photon cross section measurement, described later in Sec. 3.7, is compared to the theoretical predictions computed with fixed-order NLO and NNLO calculations, obtained with parton-level MC generators (DIPHOX [134]+GAMMA2MC [142] and 2γNNLO [143]), and with the truth-particle-level di-photon spectra predicted by leading-order parton-shower MC generators used in the ATLAS full simulation (PYTHIA and SHERPA). The contribution from the decays of a SM Higgs boson of mass 125 GeV to di-photons is not included in the predictions, as it is expected to be around 1% of the signal in the $120 < m_{\gamma\gamma} < 130$ GeV interval and negligible elsewhere. The contribution from multiple parton interactions is also estimated to be negligible and not further considered.

The main differences between the four predictions are the following:

- 2γNNLO provides a NNLO calculation of the direct part of the di-photon production cross section (including the box diagram), but neglects the fragmentation component, whose contribution, after the isolation requirement, is expected to be significantly suppressed. Technically 2γNNLO implements a different isolation requirement than the fixed-cone one, based on the proposal in [144] of a smooth-cone isolation procedure ("Frixione isolation"); using DIPHOX which implements both the fixed-cone and smooth-cone isolation prescriptions, the authors of 2γNNLO estimate the bias on the theoretical cross-section from the missing fragmentation component and the use of the Frixione isolation to be of the order of just a few %.
- DIPHOX provides a NLO calculation of both the direct and the fragmentation parts of the di-photon production cross section. It also includes the contribution from the box diagram ($gg \to \gamma\gamma$), which is in principle a term of the NNLO expansion in the strong coupling constant α_s, but – due to the large gluon luminosity at the LHC – gives a contribution com-

parable to that of the LO terms. Higher-order contributions to the gg amplitude, technically at NNNLO but of size similar to that of NLO terms, are added to the Diphox calculation using the gamma2mc program.

- Pythia provides LO matrix elements for di-photon production (Born and box diagrams) and models the higher-order terms through γ-jet and di-jet production in combination with initial-state and/or final-state radiation. It also features parton showering and an underlying-event model;
- Sherpa has features similar to those of Pythia, and in addition includes the di-photon higher-order real-emission matrix elements. For the study presented here, up to two additional partons are generated, and the photon is required to have an angular separation $\Delta R > 0.3$ from the partons from the matrix element.

The nominal factorization (μ_F), renormalization (μ_R), and – in the case of Diphox and gamma2mc – fragmentation (μ_f) scales are set in all cases to the di-photon invariant mass, $m_{\gamma\gamma}$. Different PDF sets are used by each program: CT10 NLO for Diphox and gamma2mc, MSTW2008 NNLO [145] for 2γNNLO, CTEQ6L1 [140] for Sherpa and MRST2007 LO* [146] for Pythia. Similarly to the case of photon-jet events described in the previous section, to account for non-perturbative effects not included in fixed-order parton-level generators, simulated di-photon events are produced with various parton-shower Monte Carlo generators to evaluate the ratio of generator-level cross sections with and without hadronization and underlying event, and subsequently, the parton-level cross sections are multiplied bin-by-bin by this ratio. The theory uncertainties for the NLO and NNLO predictions include statistical uncertainties, scale uncertainties, as well as PDF and non-perturbative-correction uncertainties computed in the same way as for the photon-jet cross section. For Diphox and gamma2mc, scale uncertainties are evaluated by varying each scale between $m_{\gamma\gamma}/2$ and $2m_{\gamma\gamma}$, and the envelope of all variations is taken as a systematic error; the final uncertainty is dominated by the configurations in which the scales are varied incoherently. For 2γNNLO, the scale uncertainty is evaluated by considering the variation of the predicted cross sections in the two cases $\{\mu_R = m_{\gamma\gamma}/2, \mu_F = 2m_{\gamma\gamma}\}$ and $\{\mu_R = 2m_{\gamma\gamma}, \mu_F = m_{\gamma\gamma}/2\}$.

Both Pythia and Sherpa are expected to underestimate the total cross section, because of the missing NLO (and higher-order) contributions. In the regions of low $p_{\mathrm{T},\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma} \approx \pi$, where multiple soft gluon emission is important, fixed-order calculations are expected to exhibit infrared divergences, while Pythia and Sherpa are expected to better describe the shape of the differential distributions, thanks to the effective all-order resummation of the leading logs performed by the parton shower. Finally, 2γNNLO is expected to underestimate the data in regions populated by the contribution from fragmentation (low $\Delta\phi_{\gamma\gamma}$ and $m_{\gamma\gamma}$, and $\cos\theta^*_{\gamma\gamma} \approx \pm 1$).

3.3 Measurement of the inclusive production cross section of isolated prompt photons at 7 TeV

In this section I will summarize two published measurements of the inclusive isolated prompt photon production cross section, $pp \to \gamma + X$, as a function of the photon transverse energy, using pp collision data collected in 2010 with the ATLAS detector at the LHC at a center-of-mass energy of 7 TeV. The former [14] is based on an integrated luminosity $\int Ldt = (0.88 \pm 0.1)$ pb^{-1}, and provides a measurement of the cross section for $15 \leq E_T^\gamma < 100$ GeV in the photon pseudorapidity intervals $[0, 0.6)$, $[0.6, 1.37)$ and $[1.52, 1.81)$. It was one of the very first ATLAS publications based on multi-TeV pp collisions, and laid the basis for the following photon-related ATLAS measurements. The latter [15] uses the full 2010 data sample, corresponding to $\int Ldt = (34.6 \pm 1.2)$ pb^{-1}, covers the transverse energy range $45 \leq E_T^\gamma < 400$ GeV and explores an additional pseudorapidity interval, $[1.81, 2.37)$. The corresponding kinematic region of $x_\mathrm{T} = 2E_\mathrm{T}^\gamma/\sqrt{s}$ (equal to the incoming parton fractional momentum in the leading-order approximation for photons produced at central pseudorapidity, $\eta^\gamma = 0$) is $0.004 < x_\mathrm{T} < 0.114$, while 225 GeV$^2 \leq Q^2 \leq 1.6 \times 10^5$ GeV^2.

3.3.1 Event selection

Events are triggered using a single-photon high-level trigger (Sec. 1.6.6) with a nominal transverse energy threshold of 10 GeV [14] or 40 GeV [15], seeded by L1 triggers with nominal thresholds of 5 GeV or 30 GeV, respectively. Events in which the calorimeters or the inner detector are not fully operational, or show data quality problems, are discarded. To suppress non-collision backgrounds, events are required to have at least one reconstructed primary vertex, with at least three associated tracks, consistent with the average beam spot. The remaining amount of non-collision background is estimated to be negligible using control samples collected – during normal data-taking conditions – with dedicated, low-threshold triggers in events where either no proton bunch or only one of the two beams crosses the interaction region.

Photons are reconstructed from electromagnetic clusters and tracking information provided by the inner detector as described in Sec. 1.6.1. Photon candidates near regions of the calorimeter affected by read-out or high-voltage failures are not considered. Events are selected if they contain one photon candidate in the nominal acceptance:

- $E_{\mathrm{T}}^{\gamma} > 15$ GeV, $|\eta^{\gamma}| < 1.37$ or $1.52 \leq |\eta^{\gamma}| < 1.81$ in [14],
- $E_{\mathrm{T}}^{\gamma} > 45$ GeV, $|\eta^{\gamma}| < 1.37$ or $1.52 \leq |\eta^{\gamma}| < 2.37$ in [15].

Background from jets faking photons is suppressed using shower-shape and isolation variables. Photon candidates are required to pass the tight identification criteria, based on nine discriminating variables computed from the lateral and longitudinal profiles of the energy deposited in the calorimeters, as described in Sec. 1.6.4. The photon isolation transverse energy $E_{\mathrm{T}}^{\mathrm{iso}}$ in a cone of radius 0.4 in the $\eta - \phi$ plane around the photon axis, computed as described in Sec. 1.6.5, is required to be less than 3 GeV: this criterion is expected to reject roughly 50% of background candidates with transverse energy greater than 15 GeV. The final sample size is 110 thousand events in [14], and 173 thousand events in [15]; about 30% of the photon candidates are reconstructed from conversions. The transverse energy distribution of the selected candidates is shown in Fig. 3.5.

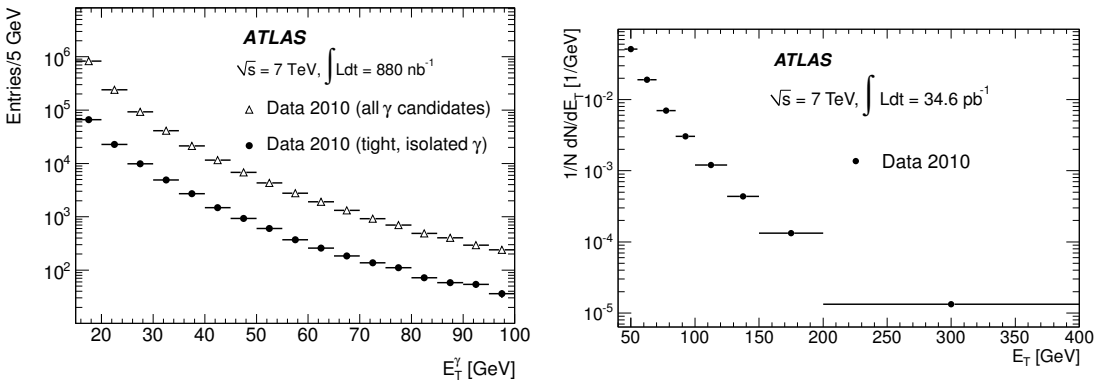

Figure 3.5: Left: transverse energy distribution of photon candidates selected in 0.88 pb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV, before (open triangles) or after (full circles) requiring tight identification criteria and isolation transverse energy lower than 3 GeV [14]. Right: transverse energy distribution of photon candidates in 35 pb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV, after the final selection [15].

3.3.2 Background subtraction

The main source of background in the selected sample is due to misidentified QCD jets, containing a neutral meson (π^0, η) that carries most of the jet energy and decays to a collimated photon pair. The background yield is estimated and then subtracted by means of a data-driven counting technique ("two-dimensional sideband method") based on the observed number of events in the control regions of a two-dimensional plane (Fig. 3.6) formed by the photon isolation transverse energy and a photon identification variable, exploiting two properties: the negligible correlation

between these two variables for background events and the dominance of background over signal in the three background control regions. To this purpose, the photon candidates are classified as:

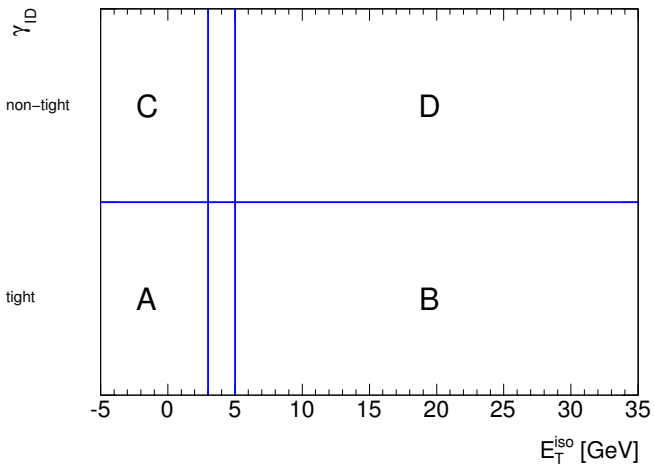

Figure 3.6: Illustration of the two-dimensional plane, defined by means of the isolation transverse energy and a subset of the photon identification (ID) variables, used for estimating, from the observed yields in the three control regions (B, C, D), the background yield in the signal region (A).

- Isolated, if $E_{\mathrm{T}}^{\mathrm{iso}} < 3$ GeV;
- Non-isolated, if $E_{\mathrm{T}}^{\mathrm{iso}} > 5$ GeV;
- Tight, if they pass the tight photon identification criteria;
- Non-tight, if they fail at least one of the tight requirements on four shower-shape variables (ΔE, E_{ratio}, F_{side}, $w_{s\,3}$) computed from the energy deposits in a few cells of the first layer of the electromagnetic calorimeter, but pass all the other tight identification criteria.

In the plane formed by the photon isolation transverse energy and the photon tight identification variable, we define four regions:

- A: the *signal* region, containing tight, isolated photon candidates.
- B: the *non-isolated* background control region, containing tight, non-isolated photon candidates.
- C: the *non-identified* background control region, containing isolated, non-tight photon candidates.
- D: the background control region containing non-isolated, non-tight photon candidates.

The number of signal events N_A^{sig} in the selected data sample is given by $N_A^{\mathrm{sig}} = N_A - (N_B - c_B N_A^{\mathrm{sig}}) \frac{(N_C - c_C N_A^{\mathrm{sig}})}{(N_D - c_D N_A^{\mathrm{sig}})}$, where N_A is the total number of events in the selected sample, N_K (for $K \in \{B, C, D\}$) are the number of events in the three control regions and $c_K \equiv N_K^{\mathrm{sig}}/N_A^{\mathrm{sig}}$ are *signal leakage fractions*, extracted from simulated signal events (their size does not exceed a few %).[2] The procedure is applied separately for each of the pseudorapidity intervals under study and in several bins of photon transverse energy. The characteristics of signal events, including the values of the leakage fractions c_K in the various η and E_{T} intervals, are studied with MC samples of photon-jet events generated using Pythia 6.421 with the modified leading-order MRST2007

2. To find the number of signal events in region A one thus have to solve a simple 2nd-order polynomial equation in N_A^{sig} and to propagate the uncertainties on the observed yields in data and signal Monte Carlo. Note that this formula is a simple extension, to the case of non-negligible signal leakage in the control regions, of the trivial relation $N_A^{\mathrm{sig}} = N_A - N_B \frac{N_C}{N_D}$ that would hold if signal events populated only region A. Note also that the only inputs from the simulation are ratios c_K of signal yields.

parton distribution functions. For the study of systematic uncertainties related to the choice of the event generator and the parton-shower model, alternative samples are generated with HERWIG 6.5 [147], which also uses LO pQCD matrix elements, but has a different parton-shower model (angle-ordered instead of p_{T}-ordered), a different hadronization model (the cluster model) and a different underlying-event model, generated using the JIMMY package [148] with multiple-parton interactions enabled. The correlations between the photon identification and isolation variables in background events are confirmed to be small using simulated QCD (di-jet) background events and inspecting the data in the background-dominated region of $E_{\mathrm{T}}^{\mathrm{iso}} > 7$ GeV, as illustrated also by the good agreement of the isolation distributions of tight and non-tight photon candidates in data in that region (Fig. 3.7).

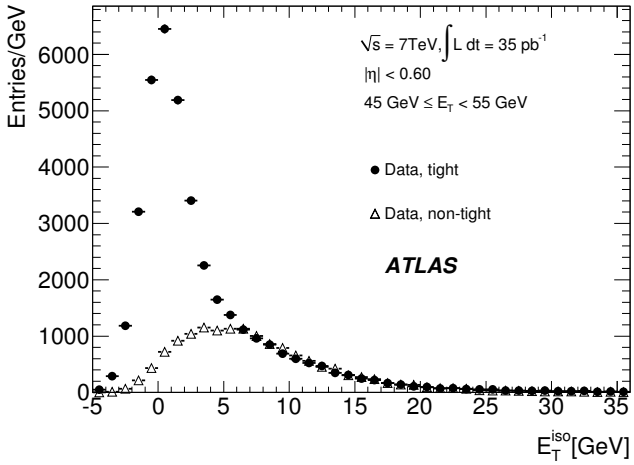

Figure 3.7: Distributions of $E_{\mathrm{T}}^{\mathrm{iso}}$ for photon candidates with 45 GeV $< E_{\mathrm{T}} <$ 55 GeV in $|\eta| < 0.6$ passing the tight (solid dots) and non-tight (open triangles) shower-shape-based selection criteria. The non-tight distribution is normalized to the tight distribution for $E_{\mathrm{T}}^{\mathrm{iso}} > 5$ GeV (non-isolated region), where the signal contamination is fairly small.

The (small) background contribution from isolated electrons from W and Z is estimated from simulated W/Z events, using the $e \to \gamma$ fake rate ($f_{e\to\gamma} \approx 8\%$) measured in data using $Z \to ee$ decays, $f_{e\to\gamma} = N_{\gamma e}^{Z}/(2N_{ee}^{Z})$, where $N_{\gamma e}^{Z}$ and N_{ee}^{Z} are the numbers of γe and ee pairs with invariant mass within 1.5 σ of the Z boson mass. The numbers of continuum background events are estimated from the sidebands of the ee and γe invariant mass distributions ($51-61$ GeV and $121-131$ GeV), and subtracted from N_{ee}^{Z} and $N_{\gamma e}^{Z}$, respectively.

The estimated signal purity N_A^{sig}/N_A increases from around 50% at $E_{\mathrm{T}} = 15$ GeV to 90% and above for $E_{\mathrm{T}} \geq 100$ GeV (Fig. 3.8). The results are cross-checked with a binned maximum likelihood fit to the isolation distribution of photon candidates passing the tight identification criteria, after relaxing the isolation requirement. The distribution is fit with the sum of a signal template and a background template, determined from control samples extracted from data, and whose normalizations are floated in the fit. The signal template is determined from the $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution of electrons from W and Z decays, shifted (100-600 MeV depending on $|\eta|$) to compensate for the differences between electrons and photons observed in the simulation. Electrons from W decays are required to pass tight electron identification criteria (Sec. 1.5), they must be accompanied by $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV, and the electron-$E_{\mathrm{T}}^{\mathrm{miss}}$ system must have a transverse mass larger than 40 GeV. Electrons from Z decays are selected with looser criteria, but the ee pair must have an invariant mass close to the Z mass. The background template is extracted for each (E_{T}, $|\eta|$) bin from non-tight photon candidates selected in data. A simulation-based correction, typically of the order of 3-4%, is applied to the final photon fraction to account for signal which leaks into the background template. The results are in good agreement with those from the counting method.

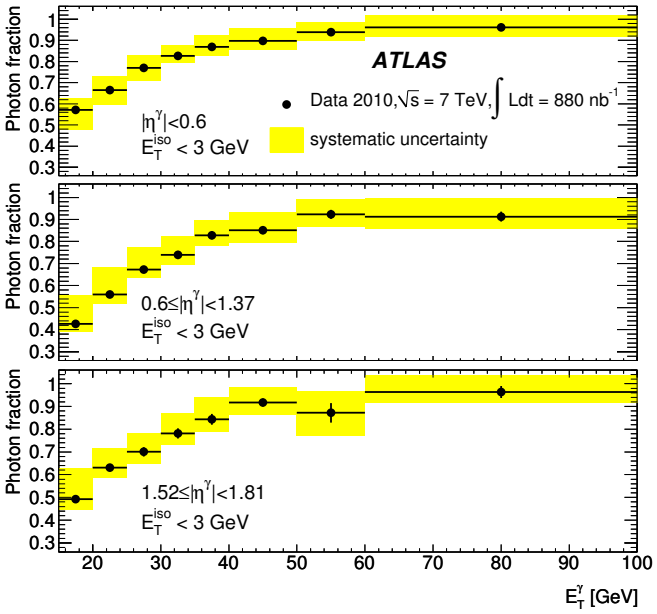

Figure 3.8: Fraction of isolated prompt photons as a function of transverse energy in the photon candidate sample selected in 0.88 pb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV, for the three pseudorapidity ranges under study.

3.3.3 Cross section measurement

The E_{T}-differential cross section for each pseudorapidity interval is determined from the estimated signal yield as a function of the photon transverse energy and the photon trigger ($\varepsilon_{\mathrm{trig}}$), reconstruction ($\varepsilon_{\mathrm{rec}}$) and identification ($\varepsilon_{\mathrm{ID}}$) efficiencies:

$$\frac{d\sigma}{dE_{\mathrm{T}}^{\gamma}} = \frac{N_A^{\mathrm{sig}}\, U}{\left(\int Ldt\right)\ \Delta E_{\mathrm{T}}^{\gamma}\ \varepsilon_{\mathrm{trig}}\ \varepsilon_{\mathrm{rec}}\ \varepsilon_{\mathrm{ID}}}. \tag{3.2}$$

The trigger efficiency, relative to reconstructed photons passing the isolation and identification requirements, is measured in data using a bootstrap procedure based on unbiased or lower-threshold triggers, and is found to be close to 100% with a relative uncertainty around 0.5%. The reconstruction efficiency $\varepsilon_{\mathrm{rec}}$ on the other hand is determined from simulated signal samples and found to be around 82%, largely due to the inefficiency (10%) introduced by the dead readout regions (that have been fixed after the end of the 2010 data-taking). It is computed as the probability of a prompt photon, with generator-level E_{T} and η in the nominal acceptance and with truth-particle-level isolation transverse energy, in a cone of radius 0.4 around the photon direction, lower than 4 GeV, to be reconstructed by the ATLAS algorithms and to pass the experimental E_{T}, η and $E_{\mathrm{T}}^{\mathrm{iso}}$ requirements. Due to the small size of the 2010 data sample, which does not allow a precise *in situ* measurement of $\varepsilon_{\mathrm{ID}}$, also the photon identification efficiency is determined from the simulation, after shifting the values of the shower-shape variables in order to reproduce their distributions in data, as described in Sec. 2.1.1. A less precise data-driven measurement of $\varepsilon_{\mathrm{ID}}$ using electrons from W decays (selected based on the track-fit quality, the presence of transition radiation in the TRT, large $E_{\mathrm{T}}^{\mathrm{miss}}$ and small electron isolation energy) is in agreement with the nominal result. $\varepsilon_{\mathrm{ID}}$ increases from around 65% to $\approx 95\%$ as a function of E_{T}^{γ}. Correction factors U (close to 1) take into account migrations between neighbouring bins in E_{T} due to energy resolution; they are obtained with various *unfolding* techniques, using the $E_{\mathrm{T}}^{\mathrm{true}} \leftrightarrow E_{\mathrm{T}}^{\mathrm{rec}}$ response matrix from simulated true photons. The simplest method is a bin-by-bin unfolding, which computes the factor U for bin i as the ratio, in the signal simulation, of reconstructed events with reconstructed E_{T} in bin i to the number of reconstructed events having true E_{T} in bin i. More sophisticated methods which better account for migrations between bins are based on the repeated (iterative) application of

Bayes' theorem [149] or on a regularization of the inverse of the response matrix [150]. Due to the excellent photon energy resolution the migrations are small and the differences between the cross sections obtained with the different methods are at the few % level and within the statistical error of the methods.

Several sources of systematic uncertainties are evaluated and their contributions are combined, taking into account their correlations. The uncertainty on the reconstruction efficiency is dominated by the following contributions:

- the uncertainty on the efficiency of the isolation selection ($3-4\%$), estimated by varying the value of the isolation criterion by the average difference (of the order of 500 MeV) observed between data and simulation for electrons from W and Z decays.
- the uncertainty from the signal-event generator kinematics and underlying-event and hadronization model (2%), obtained from the difference of the efficiencies estimated with the PYTHIA and HERWIG simulations
- the uncertainty from the limited knowledge of the detector material ($1-2\%$), evaluated by comparing the nominal simulation to an alternative one (described in Sec. 2.1.1) estimated to represent a conservative upper limit of the additional detector material that is not accounted for by the nominal simulation.

The uncertainty on the identification efficiency is dominated by the detector material uncertainty (up to 6% at low E_{T}), evaluated using the alternative detector model described previously, and the data/simulation shower-shape agreement (up to 5%), evaluated with a closure test of the shower-shape correction procedure and by using different control samples to extract the corrections, as described in Sec. 2.1.1. The uncertainty on the background-subtracted yields is dominated by two sources:

- the uncertainty on the inputs from the simulation (up to 10%), which includes the limited Monte Carlo statistics but also the variations on the leakage fractions c_K estimated with different event generators and different detector models and varying the fraction of bremsstrahlung photons in the conservative range $0-100\%$.
- the choice of the background control regions (up to 6%), estimated by varying the definition of the non-isolated and non-tight control regions.

The uncertainty on the photon energy scale ($1.5-3\%$ in 2010) translates into a global $5-10\%$ uncertainty on the cross section; the luminosity measurement introduces an additional 11% (3.5%) uncertainty on the cross section measured with 0.88 (35) pb^{-1} of data. Given the low level of pileup during 2010, the uncertainty from a possible mismodeling of the pileup dependence of the photon reconstruction, identification and isolation in the simulation is estimated to be small compared to the other sources of uncertainties.

3.3.4 Results

The measured cross sections are shown in Fig. 3.9. The red triangles represent the experimental results from [14], the black dots those from [15]. The theoretical pQCD cross sections, computed with JETPHOX with the CTEQ 6.6 PDFs and the scales set to E_{T}^{γ}, are overlaid (blue bands). In the theoretical calculation, the parton transverse energy in a cone of radius 0.4 around the photon is required to be below 4 GeV, in order to match the requirement applied in the signal MC samples to the truth-particle-level isolation of the prompt photons; varying this requirement by ± 2 GeV changes the cross section by $\pm 2\%$. PDF uncertainties lead to a systematic uncertainty decreasing from 5% to 2% with E_{T}, and varying the scales independently between 0.5 E_{T} and 2 E_{T} leads to an uncertainty decreasing from 20% to 8%. The non-perturbative corrections are consistent with unity within 2%.

The observed cross sections rapidly decrease as a function of the increasing photon transverse energy, spanning almost six orders of magnitude. The precision of the measurement is limited by its systematic uncertainty, which receives important contributions from the energy scale uncertainty, the luminosity, the photon identification efficiency, and the uncertainty on the residual background contamination in the selected photon sample. The measured prompt photon production cross

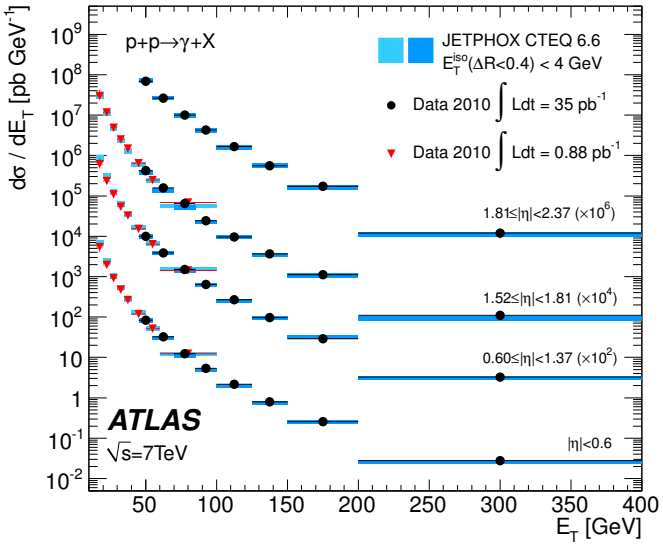

Figure 3.9: Measured (dots) and expected (shaded area) inclusive prompt photon production cross sections, as a function of the photon E_{T} and in the range $|\eta| < 0.6$, $0.6 \leq |\eta| < 1.37$, $1.52 \leq |\eta| < 1.81$ and $1.81 \leq |\eta| < 2.37$. The CTEQ 6.6 PDF is used in the JETPHOX theoretical computation (the full theoretical error is shown).

section is more than a factor of thirty larger than that measured at the Tevatron [151, 152], and a factor of 10^4 larger than for photoproduction at HERA [153, 154], assuming a similar kinematic range in transverse energy and pseudorapidity.

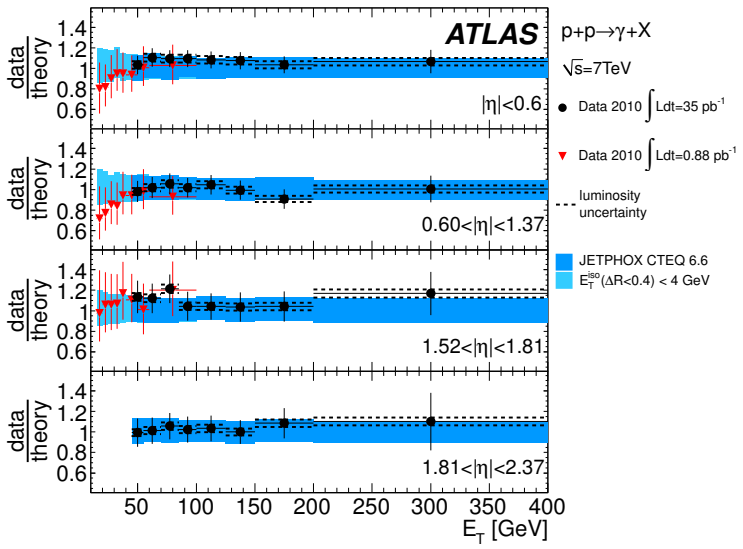

Figure 3.10: Ratio between the measured and expected inclusive prompt photon production cross sections (dots), as a function of the photon E_{T} and (from top to bottom) in the range $|\eta| < 0.6$, $0.6 \leq |\eta| < 1.37$, $1.52 \leq |\eta| < 1.81$ and $1.81 \leq |\eta| < 2.37$. The CTEQ 6.6 PDFs are used in the JETPHOX theoretical computation. The shaded area corresponds to the full theoretical error.

The two measurements are in good agreement in the overlapping region $45 < E_{\mathrm{T}} < 100$ GeV, $|\eta| < 1.81$. They are in agreement with the theoretical predictions by JETPHOX for $E_{\mathrm{T}}^{\gamma} \gtrsim 35$ GeV;

for lower E_{T}, which correspond to rather small values of x_{T} ($< 1\%$) and where the contribution from parton-to-photon fragmentation is larger, the theory tends to overestimate the data, possibly hinting either to a slight overestimation of the fragmentation functions at low transverse momentum or to an underestimation of the missing higher-order corrections (with the simple procedure of varying the energy scales in the fixed-order calculation by a factor 2 around the nominal value) and thus to the need of more accurate (NNLO) predictions. It is interesting to note that CMS observes a similar deviation in their measurement of the inclusive prompt photon production cross section at low E_{T}, reported in Ref. [155]. Recently ATLAS has presented a preliminary measurement of the inclusive isolated prompt photon production cross section up to 1 TeV [127]; the results are in agreement within uncertainties with the NLO JETPHOX prediction over the whole range of 100 GeV $< E_{\mathrm{T}} <$ 1 TeV.

3.4 Impact of the isolated photon data on the gluon distribution in the proton

The last PDF analysis to include prompt photon data was done more than 10 years ago for the MRST99 parametrization [156]. Since then, although the number of available prompt photon measurements at colliders has continuously increased, these and other results from fixed-target experiments, spanning center-of-mass energies of $\sqrt{s} \approx 20 - 7000$ GeV, have not been used in PDF fits. This choice was motivated by the discrepancy between the results on the inclusive (non-isolated) prompt photon production cross section from a single fixed target experiment at $\sqrt{s} \approx 30$ GeV, E706 [157], in 1998, and the theoretical predictions available at that time. However, as noted in [135] by the authors of JETPHOX, who were able around 2002 to extend the NLO pQCD calculation of prompt photon cross sections to the case of isolated photons, if one considers the experimental data for isolated photons at larger center-of-mass energies, where the contribution from parton-to-photon fragmentation and the non-perturbative effects are significantly reduced, a very good data-NLO pQCD agreement is observed over nine orders of magnitudes in the cross section and two orders of magnitude in x_{T}, as illustrated in Fig. 3.11. It is therefore reason-

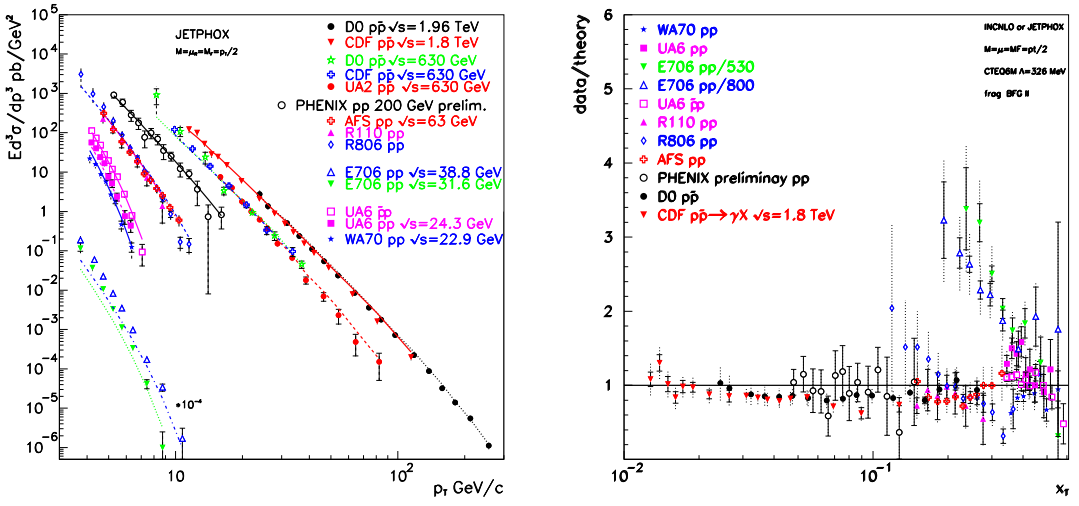

Figure 3.11: Left: pre-LHC inclusive and isolated prompt photon productions cross sections vs photon E_{T}, measured in pp and $p\bar{p}$ collisions compared to JETPHOX NLO predictions using BFG-II (CTEQ6M) for fragmentation (structure) functions and a common scale $E_{\mathrm{T}}/2$. For clarity's sake the E706 data are scaled by a factor 10^{-4}. Right: ratios between data and theory, as a function of x_{T}. For PHENIX and lower-energy data the inclusive cross section is used while the isolated one is used for CDF and D0 [135].

able to re-include in PDF fits the available isolated photon data. This is also supported by the

observation [158] that, after removing the E706 and other low-energy, non-isolated cross section measurements, all the experimental results plotted as a function of x_T follow clear power-law dependencies and coalesce over a single curve when normalised by $\sqrt{s}^n$ with $n \approx 4.5$, which is very close to the $1/p_T^4$ dependence expected for partonic $2 \to 2$ scattering cross sections in the conformal QCD limit. This is also the case when including the results (as well as those by CMS) presented

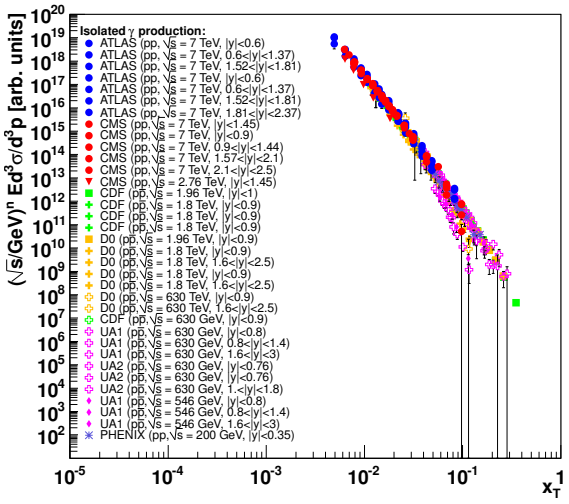

Figure 3.12: Isolated-photon spectra measured in pp and $p\bar{p}$ collisions at collider energies as a function of x_T, where the invariant cross sections have been. scaled by $\sqrt{s}^{4.5}$. [158].

here (see Fig. 3.12), and is a strong indication of the perturbative origin for the production of isolated photons, thus justifying the use of pQCD calculations. A global PDF analysis that includes the isolated photon data and whose results are compared to the JETPHOX predictions in order to extract constraints on the gluon distribution function of the proton has indeed been made [158], using the following NNPDF Bayesian reweighting technique [159]:

- the data (measurements and error covariance matrices) in the standard NNPDF global fit are used to generate $N = 100$ Monte Carlo datasets which are then fitted to determine N replicas f_k of the PDFs, each one with equal weight; the expected value for any observable $X[f]$ is then computed as $\langle X \rangle = \frac{1}{N} \sum_{k=1}^{N} X[f_k]$.
- when a new independent dataset (corresponding to the measurements $\{y_1..y_n\}$ and their error matrix σ_{ij}, $i,j = 1..n$) is added, the chi-squared χ^2_k between each PDF replica f_k and the data is calculated, $\chi^2_k = \frac{1}{n} \sum_{i,j=1}^{n} (y_i - y_i[f_k]) \sigma_{ij}^{-1} (y_j - y_j[f_k])$, assuming Gaussian uncertainties.
- based on Bayes' theorem, each replica is assigned a new weight, $w_k = P(f_k|\chi^2_k) \propto P(\chi^2_k|f_k) = (\chi^2_k)^{n/2-1} e^{-\frac{1}{2}\chi^2_k}$, and the new central value of X is $\langle X \rangle = \frac{1}{N} \sum_{k=1}^{N} w_k X[f_k]$.
- the effective number of replicas after reweighting is quantified using Shannon's entropy, $N_{\text{eff}} = \exp\{\frac{1}{N} \sum_{k=1}^{N} w_k \ln(N/w_k)\}$.

Since the correlations between the differential cross section measurements $d\sigma/dE_T^\gamma$ in different E_T^γ bins were not provided by the experiments, they are neglected in the χ^2 calculation. The main results are the following:

- a global χ^2 of 1.1 is found for the agreement between data and theory.
- considering $Q = 100$ GeV as a benchmark, only the LHC data lead to a significant uncertainty reduction, up to 20%, on $g(x, Q^2)$ at $x \approx 0.002 - 0.05$ (Fig. 3.13), while the Tevatron data bring smaller improvements at $x \approx 0.01 - 0.02$ and the other datasets have negligible impact.
- the central value of the gluon PDF is essentially unaffected, indicating good agreement between the LHC photon results and the other gluon PDF determinations (Tevatron jet data and DIS data from Hera).

- for three different scales $Q = 3.16, 10$ and 100 GeV, the maximum $g(x, Q^2)$ uncertainty reduction is $10 - 20\%$ in the region $x \approx 0.01 - 0.05$.
- the constraints on the quark PDFs are negligible.
- the PDF uncertainty on the gluon-fusion Higgs boson production cross section, which constitutes $\approx 87\%$ of the total production cross section, can be reduced by 20% for a light boson with a mass of 120 GeV, if photon data are used (Fig. 3.13). Since the theoretical uncertainties on the Higgs boson production cross section and branching ratios dominate the systematic uncertainties on the current measurements, this reduction is rather important in perspective for improving the accuracy of those results.

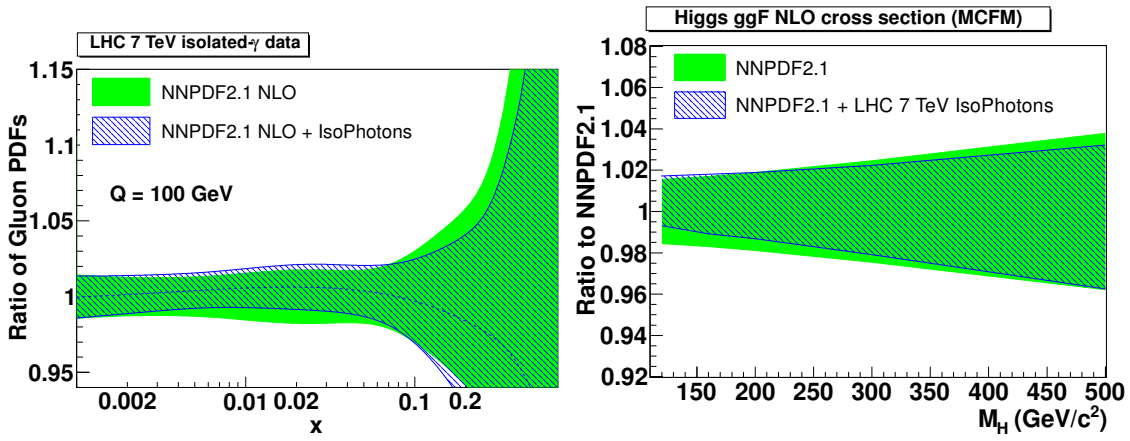

Figure 3.13: Left: comparison between the NNPDF2.1 NLO gluon before (green solid band) and after (dashed blue area) inclusion of the LHC isolated prompt photon cross section data, with PDFs evaluated at $Q = 100$ GeV. Right: ratio of Higgs production gluon-fusion cross sections with NNPDF2.1 NLO PDFs before and after including LHC isolated photon data [158].

It should be noted that the JETPHOX scale uncertainties on the theoretical cross section are not included in the χ^2 calculation since there is no consensus yet among the PDF fitting groups on their statistical treatment. Indeed, all PDF global fits neglect scale uncertainties on the predicted cross sections that are compared to the data to extract the PDFs; this should have a minor effect for the majority of the observables, that are compared to NNLO calculations with small scale uncertainties, but can be relevant for photon and jet observables that have to be compared to NLO calculations.

3.5 Measurement of the production cross section of isolated prompt photons in association with jets at 7 TeV

The differential cross section $d\sigma/dE_{\mathrm{T}}^{\gamma}$ for the process $pp \to \gamma + jet + X$ has been measured for isolated photons in the pseudorapidity range $|\eta^{\gamma}| < 1.37$ and transverse energy $E_{\mathrm{T}}^{\gamma} > 25$ GeV, produced in association with a high-p_{T} (> 20 GeV) jet, using the full pp collision data set collected by ATLAS in 2010 [16].

Events are recorded using two single-photon triggers, with nominal transverse energy thresholds of 20 and 40 GeV. The former, used to collect events in which the photon transverse energy is lower than 45 GeV, has an average prescale of 5.5, leading to a total integrated luminosity of 6.7 pb^{-1}; the latter, used to collect events in which the photon transverse energy is greater than 45 GeV, is unprescaled and the corresponding total integrated luminosity of the collected sample is $\int Ldt = 37.1$ pb^{-1}.

The same photon reconstruction algorithms, quality, identification and isolation criteria used for the inclusive prompt photon measurement are applied to the photon candidates selected for this measurement. In addition, events are required to contain at least one jet with $p_{\mathrm{T}} > 20$ GeV,

$|y| < 4.4$, not overlapping with the photon or with isolated electrons ($\Delta R_{\gamma j} > 0.3$, $\Delta R_{ej} > 0.3$). Quality criteria are applied to the jet candidates to suppress fake ones from calorimeter noise, cosmic rays and beam-related backgrounds. In case of multiple photon (the average multiplicity is 1.02) or jet (average multiplicity around 1.4) candidates in the same event, that with the higher transverse momentum is retained. A minimum separation of $\Delta R > 1.0$ in the $\eta - \phi$ plane is required between the jet and the photon.

The cross sections are determined separately for the three jet rapidity intervals $|y^{\text{jet}}| < 1.2$, $1.2 \leq |y^{\text{jet}}| < 2.8$ and $2.8 \leq |y^{\text{jet}}| < 4.4$, distinguishing between the same-sign ($\eta^{\gamma} y^{\text{jet}} \geq 0$) and opposite-sign ($\eta^{\gamma} y^{\text{jet}} < 0$) configurations. This subdivision allows the comparison between data and theoretical predictions, computed with the NLO pQCD calculation by Jetphox corrected for non-perturbative effects, in configurations where the relative contribution of the fragmentation component to the total cross section is different (see Fig. 3.14), and in different ranges of x. The

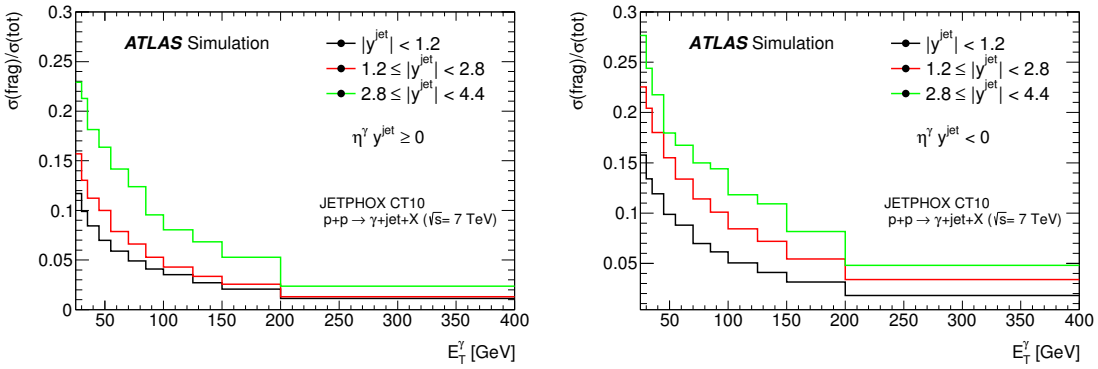

Figure 3.14: Fraction of the total photon-jet cross section due to the fragmentation process, for the same-sign (left) and opposite-sign (right) angular configurations, as computed with Jetphox. The photon is required to have $E_{\text{T}} > 25$ GeV and $|\eta| < 1.37$, and the jet has $p_{\text{T}} > 20$ GeV. [3]

differential cross sections are measured up to $E_{\text{T}}^{\gamma} = 400$ GeV for $|y^{\text{jet}}| < 2.8$ and up to $E_{\text{T}}^{\gamma} = 200$ GeV otherwise, covering the region $x \gtrsim 0.001$, 625 GeV$^2 \leq Q^2 \leq 1.6 \times 10^5$ GeV^2, thus extending the kinematic reach of previous photon + jet measurements at hadron [160, 161, 162, 163] and electron-proton [153, 164, 165, 166] colliders.

The number of events selected in the sample collected with the 20 (40) GeV trigger is 96 (117) thousand. The transverse energy distribution of the photon candidates in the selected sample is shown in Fig. 3.15.

The signal yield in the selected sample is estimated with the same counting technique exploited for the measurement of the inclusive photon production cross section. The signal purity typically increases from $50 - 70\%$ at $E_{\text{T}}^{\gamma} = 25$ GeV to above 95% for $E_{\text{T}}^{\gamma} > 150$ GeV, as shown in Fig. 3.16. The effect of the non-negligible signal leakage in the background control regions increases the measured purity by $5 - 6\%$ at $E_{\text{T}}^{\gamma} = 25$ GeV and $\approx 2\%$ at $E_{\text{T}}^{\gamma} > 150$ GeV compared to the case where the signal in the background regions is assumed to be negligible.

The combined signal trigger, reconstruction, and selection efficiency is evaluated from simulated signal samples generated with Pythia 6.423 and Herwig 6.510 with the modified leading-order MRST2007* LO PDFs. At truth-particle level the photon and the jet are required to pass the same kinematic requirements as the reconstructed objects; the photon isolation (computed from the true four-momenta of the generated particles inside a cone of radius 0.4 around the photon direction) is required to be lower than 4 GeV. The event selection efficiency typically rises from 50% to 80% as a function of E_{T}^{γ}; an inefficiency of around 15% is due to the acceptance loss originating from a few inoperative optical links in the calorimeter readout (fixed between the 2010 and 2011 data taking) and from the isolation requirement, while an inefficiency decreasing from $20 - 25\%$ to almost zero with increasing E_{T}^{γ} is due to the photon identification selection.

3. Note that this quantity is somehow unphysical as it depends on the chosen fragmentation scale and only the sum of the direct and fragmentation contributions is physically meaningful; nevertheless, the comparison between the different angular configurations shows a clear trend, according to which the fragmentation contribution is larger in opposite-sign configurations and the larger the jet rapidity.

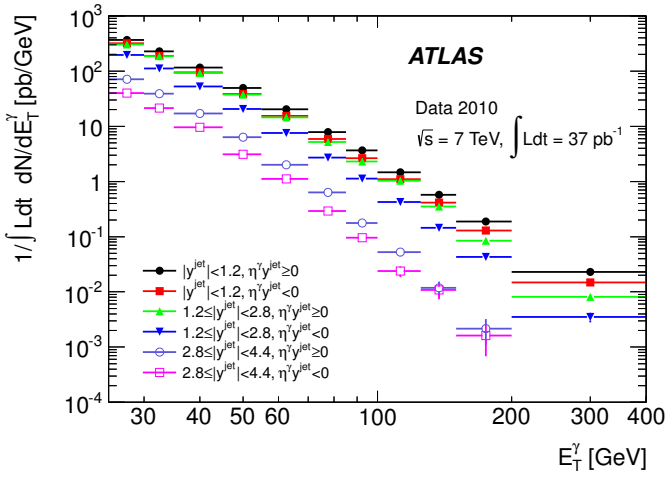

Figure 3.15: Transverse energy distribution of photon candidates in photon+jet candidate events selected in the pp collisions at $\sqrt{s} = 7$ TeV collected by ATLAS in 2010. The distribution is normalized by the integrated luminosity and the transverse-energy bin width.

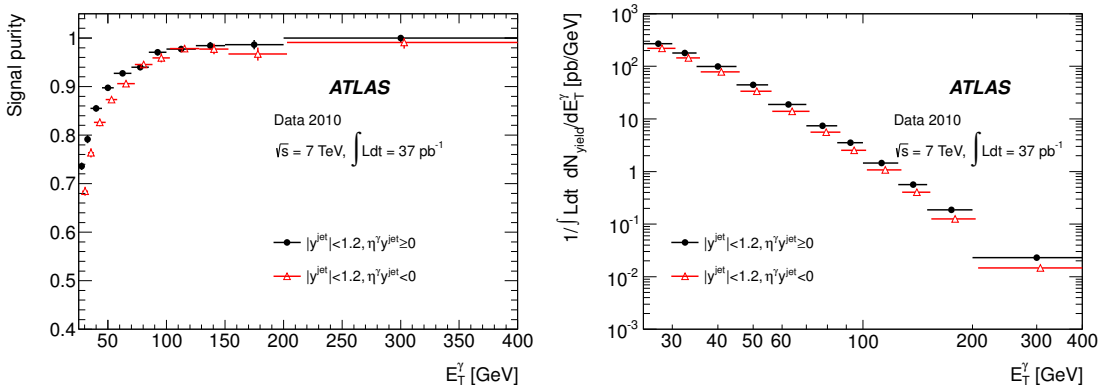

Figure 3.16: Estimated signal purity (left) and signal yield normalized by bin width and integrated luminosity (right) in data as a function of the photon transverse energy, for the same-sign (full circles) and opposite-sign (open triangles) angular configurations for events with $|y^{\mathrm{jet}}| < 1.2$. A small horizontal displacement has been added to the points corresponding to the opposite-sign configurations, so that the error bars are clearly shown. The errors are statistical only.

Systematic uncertainties on the measured cross section originate from the sources mentioned in Sec. 3.3.3 (knowledge of the detector material, event generator, underlying event and pileup model, fraction of bremsstrahlung photon-jet events, photon shower-shape corrections, photon trigger efficiency, photon energy scale and resolution, luminosity, background subtraction procedure), as well as additional uncertainties related to the jet selection. The latter include:

- jet energy scale and resolution. The cross section uncertainty is determined by varying the jet energy scales and resolutions within their uncertainties. The only non-negligible effect concerns the cross section in the first E_{T}^{γ} bin ($4-7\%$ for $|y^{\mathrm{jet}}| < 2.8$ and $9-20\%$ otherwise) and is due to the jet energy-scale uncertainty, affecting the efficiency of the $p_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV requirement.
- jet reconstruction efficiency. The ATLAS simulation is found to reproduce data jet reconstruction efficiencies to better than 2%, which is assigned as a systematic uncertainty to the measured cross section.
- simulated jet multiplicity. The LO generators used to estimate the signal efficiencies may not reproduce precisely the jet multiplicity observed in data; reweighting the simulation in order to reproduce the jet multiplicity observed in data, the cross section changes by less than 1%, which is taken as a systematic uncertainty.

The total uncertainty on the measured cross section is typically within $10-30\%$ and dominated by the systematic component, with the exception of the $1-2$ highest E_{T}^{γ} bins where the statistical uncertainty is largest and ranges between 10% and 80% depending on the y^{jet} interval.

The expected cross section is computed with Jetphox using the CT10 NLO PDFs and the energy scales set to E_{T}^{γ}. The parton transverse energy in a cone of radius 0.4 around the photon is required to be below 4 GeV. Residual non-perturbative effects are evaluated using simulated signal Pythia samples to evaluate the ratios of truth-level cross sections with and without hadronization and underlying event and to multiply each bin of the Jetphox cross sections by these ratios. The correction factors are smaller than one (around 0.9-0.95) at low E_{T}^{γ}, indicating that the impact of hadronization (which spreads energy outside of the jet area) on the jet p_{T} (and thus on the efficiency of the $p_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV cut) is more important than the extra energy added from the underlying event and pile-up. The correction factors are consistent with one for high E_{T}^{γ}, since the jet p_{T} is also large and far from the threshold and thus the $p_{\mathrm{T}} > 20$ GeV cut becomes fully efficient both at parton- and particle-level. The uncertainty $(2-4\%)$ on these corrections is estimated as the maximum spread of the correction factors obtained from Pythia using both the nominal and the Perugia 2010 tunes [47] and with Herwig++ 2.5.1 with the UE7000-2 tune [48].

As shown in Fig. 3.17 the NLO pQCD cross section calculations by Jetphox are in fair agreement with the measurements considering the typical $(10-30\%)$ experimental and theoretical systematic uncertainties in each of the six angular configurations under study, except for the $E_{\mathrm{T}}^{\gamma} \lesssim 35$ GeV region, where the NLO QCD calculation overestimates the measured cross section, as in the inclusive photon cross section measurement.

For the purpose of validating the settings of some of the main leading-order parton shower MC generators used by ATLAS, the experimental cross sections have been compared with the particle-level ones predicted by Pythia and Sherpa. [4] The Sherpa predictions agrees fairly well with the data, within 20%, in all the six angular configurations under study, though the MC cross sections tends to underestimate the data ones at low E_{T}^{γ}. The Pythia predictions agree with the data within $\approx 20\%$ in the configurations where the fragmentation component is smaller ($|y^{\mathrm{jet}}| < 1.2$, both same- and opposite-sign, or $1.2 < |y^{\mathrm{jet}}| < 2.8$, same-sign), while in the other configurations the agreement is poorer, and the Pythia cross section underestimates the data by 20-50%; at low E_{T}^{γ}, Pythia tends to overestimate the data cross section. The CMS and D0 collaborations recently presented measurements of the photon-jet cross section [167, 168] with similar conclusions.

4. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/STDM-2011-28/

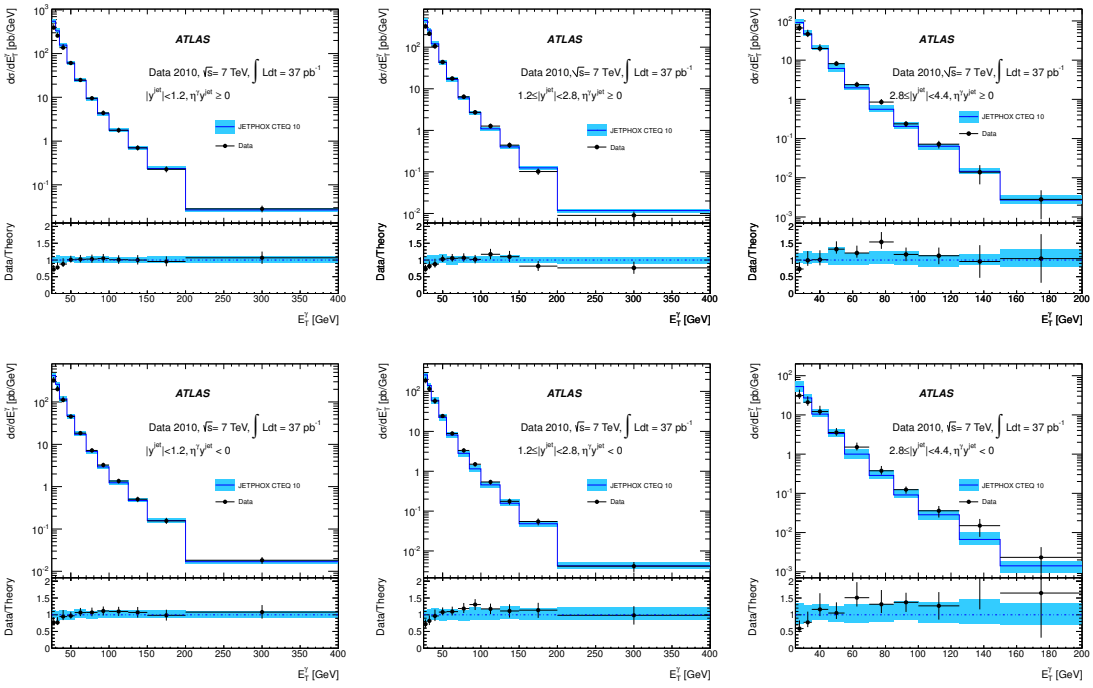

Figure 3.17: Top graphs: experimental (black dots) and theoretical (blue line) photon + jet production cross sections, for the three same-sign (top row) and the three opposite-sign (bottom row) angular configurations. The black error bars represent the total experimental uncertainty. The blue bands show the total uncertainties on the theoretical predictions obtained with JETPHOX. Bottom graphs: ratio between the measured and the predicted cross sections. The blue bands show the theoretical uncertainties while the error bars show the experimental uncertainties on the ratio. Left column: $|y^{\mathrm{jet}}| < 1.2$. Middle column: $1.2 \leq |y^{\mathrm{jet}}| < 2.8$. Right column: $2.8 \leq |y^{\mathrm{jet}}| < 4.4$.

3.6 Sensitivity of the LHC isolated γ+jet data to the proton PDFs

Using the same Bayesian reweighting technique described in Sec. 3.4, the impact of the previous results on the gluon and light quarks PDFs in the proton has been estimated [169]. This work, resulting from the collaboration of the ATLAS photon-jet analysis team with Juan Rojo and David d'Enterria, is the first one to use isolated photon-jet data at high-energy hadron collisions for such a purpose.

The potential of the measurements to constrain those PDFs has been first assessed by inspecting the correlation coefficient, as a function of x and for fixed Q^2, $E_{\rm T}^{\gamma}$ and angular configuration, between the PDF f and the cross section σ under study, $\rho = \frac{\langle f\sigma\rangle - \langle f\rangle\langle\sigma\rangle}{\sqrt{var_f var_\sigma}}$, where the averages and variances are computed over the set of the NN PDF replicas (e.g. in Fig. 3.18). The photon-jet

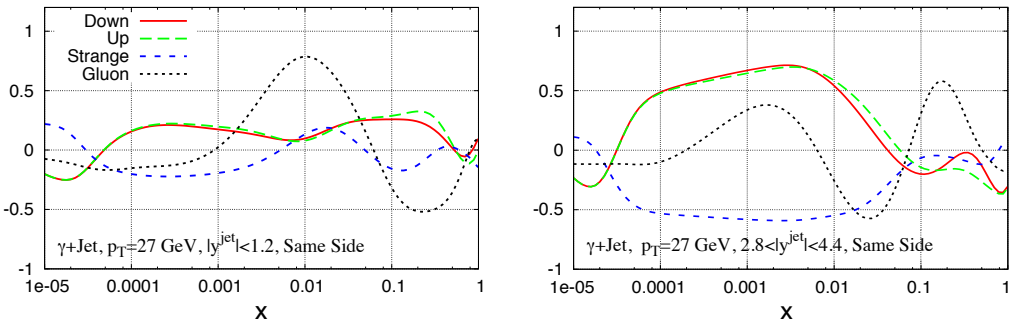

Figure 3.18: Correlations between the γ-jet cross section in pp collisions at 7 TeV and various flavours of the NNPDF2.1 parton densities for $E_{\rm T}^{\gamma} = 27.5$ GeV and central (left) or forward (right) jets in the same-sign photon-jet angular configuration.

cross sections at central jet rapidities show a dominant sensitivity to $g(x, Q^2)$ around $x = 0.01$ for low $E_{\rm T}^{\gamma}$ and around $x = 0.1$ for high $E_{\rm T}^{\gamma}$; forward-jet cross sections probe the gluon and light-quark densities for a wide range of values at medium and small x for small and moderate $E_{\rm T}^{\gamma}$, while at higher $E_{\rm T}^{\gamma}$ the light quarks PDF at very large x are probed.

Comparing the data to the Jetphox prediction, an overall good agreement over the range $x \approx 0.01 - 1$ is found, with χ^2 near one except for $2.8 < |y^{\rm jet}| < 4.4$, where $\chi^2 \approx 2 - 2.5$ is found. However, the χ^2 may be overestimated as scale uncertainties on the theory prediction, varying between 10% and 15%, have not been included in the χ^2 analysis as there is no consensus yet on how to consistently include them in global PDF analyses. The comparison of the gluon and light quark PDFs before and after adding the photon-jet data shows that the central values are essentially unaffected and the PDF uncertainties are only mildly reduced (5%) at intermediate gluon x ($\approx 0.06 - 0.3$) and small quark x ($\approx 10^{-4} - 10^{-2}$). It has also been estimated with generated pseudodata that a more refined measurement, exploiting the much larger statistics of the 2011 data to better constrain the uncertainties on the photon and jet efficiencies and calibrations, and profiting from the better knowledge of the detector material and of the integrated luminosity, would yield, thanks to an overall improvement by a factor two in the total experimental uncertainty, a reduction by up to 20% in some x regions for both the gluon and the light quark PDFs (e.g. in Fig. 3.19). These conclusions however should be revisited by propagating the Jetphox scale uncertainty to the final result, once a procedure is agreed among the PDF fitting groups on how to treat such uncertainties and once the global fits include the scale uncertainties on the predictions of the observables that are already exploited. To have an idea of the impact of the Jetphox scale uncertainty on the previous results one could redo the PDF extraction using either the nominal Jetphox scales or varying them within 0.5 and 2 times the default value and then look at the envelope of the gluon PDFs obtained in the various configurations, but it would not be straightforward to assign a statistical interpretation to such uncertainties.

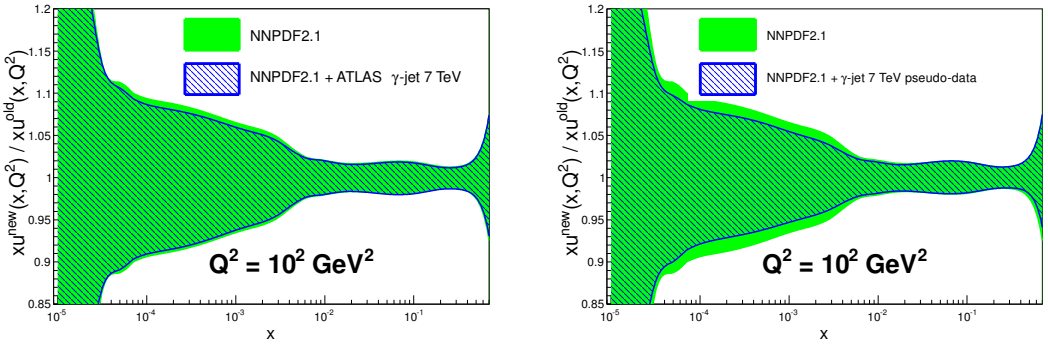

Figure 3.19: Ratio between the NNPDF2.1 NLO quark PDF and associated uncertainties before (green solid band) and after (dashed blue area) inclusion of the ATLAS γ-jet data measured at 7 TeV (left) or in the case of artificial γ-jet pseudodata at 7 TeV assuming the same kinematical distributions of the existing ATLAS measurement but with reduced experimental uncertainties. The PDFs are evaluated at $Q^2 = 100$ GeV^2.

3.7 Measurement of the production cross section of isolated photon pairs at 7 TeV

Using the *pp* collision data collected at 7 TeV, ATLAS has published two measurements of the isolated di-photon production cross section, one based on about 36 pb^{-1} of data collected in 2010 [17] and a second one based on about 4.9 fb^{-1} of data collected in 2011 [18]. Though I contributed significantly to both publications, focusing on the estimation of the background subtraction for the former one (developing a two-dimensional isolation fit method that was later used also in the $H \to \gamma\gamma$ studies) while working on almost every aspect of the analysis for the latter, I will only focus here on the most recent study, which has better precision (due to an increase in statistics by a factor larger than 100 and reduced systematic uncertainties) and provides a larger set of comparisons between data and theoretical predictions, including a NNLO calculation.

The integrated di-photon production cross section is measured, as well as the differential cross sections as a function of four kinematic variables: the di-photon invariant mass ($m_{\gamma\gamma}$), the di-photon transverse momentum ($p_{\mathrm{T},\gamma\gamma}$), the azimuthal separation between the photons in the laboratory frame ($\Delta\phi_{\gamma\gamma}$), and the cosine of the polar angle of the highest E_{T} photon in the Collins–Soper di-photon rest frame ($\cos\theta^*_{\gamma\gamma}$) [170].

Events are collected using a di-photon trigger with a nominal transverse energy threshold of 20 GeV for both photon candidates. Events are then selected offline requiring two photon candidates reconstructed in the fiducial ECAL acceptance ($|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$), passing the default quality requirements and loose identification criteria. The leading-E_{T} and sub-leading-E_{T} photon candidates are required to have $E_{\mathrm{T},1} > 25$ GeV and $E_{\mathrm{T},2} > 22$ GeV, respectively, and to have an angular separation $\Delta R > 0.4$. Both photons must pass tight identification requirements and have an isolation transverse energy in the calorimeter lower than 4 GeV. The final sample size is 166 thousand events, and the fraction of events with an additional photon pair passing all criteria except for the requirement on the two photons being the leading and sub-leading E_{T} candidates is less than 1 per 100000.

After the selection, the main background is due primarily to γ-jet and secondarily to di-jet (jj) final states, collectively called "jet background". It is estimated *in situ*, on a statistical basis, with a binned extended maximum likelihood fit to the two-dimensional distribution in data of the isolation energies of the two photon candidates and cross-checked with a counting method based on an extension of the two-dimensional sideband method described in Sec. 3.3.2. For the fit, which allows the simultaneous extraction of the numbers of true $\gamma\gamma$ signal, γj, $j\gamma$ [5] and jj background events by exploiting the different isolation distributions of prompt photons and jets, the $E_{\mathrm{T}}^{\mathrm{iso}} < 4$

5. Here and in the following, γj ($j\gamma$) denotes the events where the leading (sub-leading) candidate is a true photon, and the other candidate is a true hadronic jet.

GeV requirement is relaxed to $E_{\mathrm{T}}^{\mathrm{iso}} < 8$ GeV in order to use the "non-isolated" region $4 < E_{\mathrm{T}}^{\mathrm{iso}} < 8$ GeV to normalize the background pdfs. The correlations between the isolation transverse energies of the two candidates in di-photon, γj, and $j\gamma$ events are found to be negligible in MC samples, as a consequence of the requirement on the angular separation between the two photon candidates and the narrow energy deposit of prompt photons in the calorimeter. The $(E_{\mathrm{T},1}^{\mathrm{iso}}, E_{\mathrm{T},2}^{\mathrm{iso}})$ pdfs for each of these three event species is thus the simple product of two one-dimensional templates. For the jj component instead, large correlations (about 8%) are observed, and a two-dimensional template is used. The isolation transverse energy distributions of signal photons, separately for leading and sub-leading candidates, are obtained from a Sherpa simulated di-photon sample after applying small shifts ($120 - 160$ MeV) to account for differences between data and simulation The $E_{\mathrm{T}}^{\mathrm{iso}}$ distributions of prompt photons in γj and $j\gamma$ events are assumed to be identical to that of prompt photons in di-photon events, as verified in simulated samples. The isolation transverse energy template of the jet faking a photon in $j\gamma$ (γj) events is extracted directly from data where the leading (sub-leading) candidate passes the non-tight and the other candidate passes both the tight identification and isolation requirements. For jj events, the two-dimensional $(E_{\mathrm{T},1}^{\mathrm{iso}}, E_{\mathrm{T},2}^{\mathrm{iso}})$ template is obtained from data in which the two candidates are required to be non-tight. The jet background templates are corrected for signal leakage in the control samples, estimated from the Sherpa sample. Figure 3.20 shows the one-dimensional distributions of the isolation energies

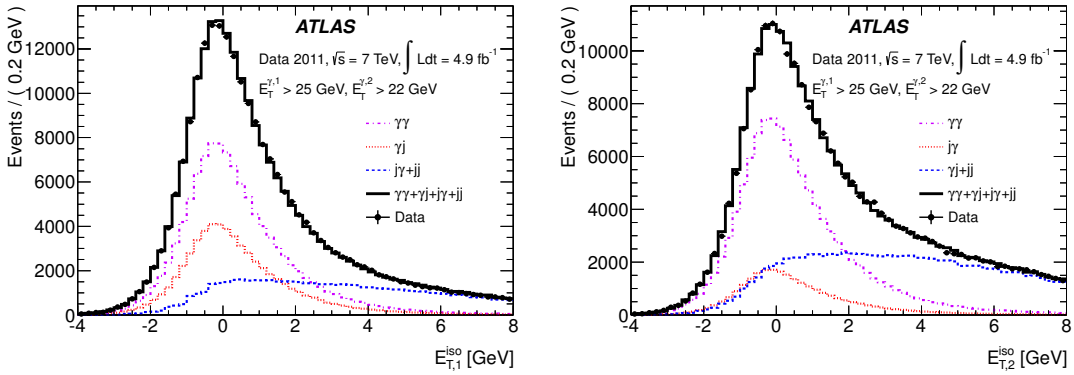

Figure 3.20: Projections of the two-dimensional fit to the isolation transverse energies of the two photon candidates: leading photon (left) and sub-leading photon (right). The photon templates from Sherpa are shifted by +160 MeV (+120 MeV) for the leading (sub-leading) photon. Solid circles represent the observed data. The (black) solid line is the fit result, the (violet) dash-dotted curve shows the $\gamma\gamma$ component. The (red) dotted line shows in the left (right) figure the contribution from γj ($j\gamma$) events. In both figures, the (blue) dashed line represents a broad background component in the photon candidates' sample: for the leading candidate this is due to $j\gamma$ and jj final states, whereas for the sub-leading candidate it comes from γj and jj final states.

of the leading and sub-leading photon candidates in data, compared to the projections of the fit. The total di-photon yield for $E_{\mathrm{T}}^{\mathrm{iso}} < 4$ GeV is estimated to be around 111 thousand, with a purity of about 67%. To obtain the differential signal yields as a function of the di-photon kinematic variables, such as $m_{\gamma\gamma}$, $p_{\mathrm{T},\gamma\gamma}$, $\Delta\phi_{\gamma\gamma}$ and $\cos\theta^*_{\gamma\gamma}$, the above methods are applied in each bin of the variable under study. A small background ($\approx 5\%$) from isolated electron misreconstructed as photons is further subtracted using the $e \leftrightarrow \gamma$ fake rates measured in data from $Z \to ee$ and $Z \to ee\gamma$ decays and the number of isolated γe and ee events reconstructed in data with similar requirements as the di-photon candidates.

Several sources of systematic uncertainty on the signal yield due to the jet background subtraction procedure are considered. The uncertainty originating from the choice of the background control regions used to extract the jet isolation distributions is estimated by varying the number of relaxed criteria in the non-tight definition. For the integrated di-photon yield, the effect is found to be $^{+3}_{-6}$%; for the differential yields, the uncertainty is at most $\pm 9\%$. The uncertainty on the photon isolation distributions, which are taken from the Sherpa di-photon sample, is evaluated by using alternative templates either from the Pythia di-photon sample, or from data. In data, the

template for the leading (sub-leading) photon is obtained – after removing the requirement $E_{\mathrm{T}}^{\mathrm{iso}} < 8$ GeV for the leading (sub-leading) photon candidate – from the difference between the isolation distribution of candidates in events where both photons pass the tight criteria and the isolation distribution (normalised to the previous one in the $7 < E_{\mathrm{T}}^{\mathrm{iso}} < 17$ GeV region) of candidates in events in which the leading (sub-leading) photon candidate fails the tight identification while the other candidate passes tight identification and isolation criteria. The PYTHIA di-photon sample exhibits higher tails (by a factor almost 2) than SHERPA at large values of $E_{\mathrm{T}}^{\mathrm{iso}}$. The data-driven template, on the other hand, is characterized by smaller tails than the SHERPA template, since it is obtained by assuming that the isolation region above 7 GeV is fully populated by background. The corresponding uncertainty on the signal yield is estimated to be $^{+2}_{-3}$% of the integrated di-photon yield, and is rather uniform as a function of $m_{\gamma\gamma}$, $p_{\mathrm{T},\gamma\gamma}$, $\Delta\phi_{\gamma\gamma}$ and $\cos\theta^*_{\gamma\gamma}$ and always below 4%, except at very low $m_{\gamma\gamma}$ where it reaches ±5%. The photon isolation template is, to a large extent, independent of the variables under study. Repeating the background subtraction procedure using photon isolation templates extracted in bins of the di-photon variable under study leads to variations of the estimated signal yield within $^{+2}_{-4}$%. Other systematic effects (differences between the photon isolation in *gamma*γ and γ+jet events, the shift of the photon isolation template, the uncertainty on the signal leakage in the non-tight control regions and the uncertainty on the detector material in the simulation) have been considered, and found to be smaller than those previously discussed.

The background-subtracted differential distributions obtained from the data are unfolded to obtain the particle-level spectra by dividing the signal yield in each bin of the di-photon observable under study by a bin-by-bin correction, which accounts for signal reconstruction and selection efficiencies and for finite resolution effects. The spectra are then divided by the trigger efficiency (measured in data with a bootstrap technique), the bin width and the integrated luminosity $\int Ldt = (4.9 \pm 0.2)$ fb^{-1}, to obtain the differential cross sections as a function of $m_{\gamma\gamma}$, $p_{\mathrm{T},\gamma\gamma}$, $\Delta\phi_{\gamma\gamma}$, and $\cos\theta^*_{\gamma\gamma}$. The bin-by-bin nominal corrections are evaluated from a SHERPA di-photon simulated sample, in which the shower-shape variables are corrected for the observed differences between data and simulation in photon-enriched control samples. Alternative corrections are calculated with the PYTHIA di-photon sample or using a simulated di-photon sample which contains additional material upstream of the calorimeter. Their effect on the total cross section is within $^{+2}_{-5}$% for $m_{\gamma\gamma}$, ±3% for $p_{\mathrm{T},\gamma\gamma}$, $^{+3}_{-4}$% for $\Delta\phi_{\gamma\gamma}$ and $^{+2}_{-3}$% for $\cos\theta^*_{\gamma\gamma}$. The effect of the uncertainty on the efficiency of the photon identification criteria is estimated by varying the identification efficiency in the simulation by its uncertainty, determined *in situ* (Sec. 2.2). The uncertainties on the electromagnetic (photon) energy scale and resolution are also propagated to the final measurement by varying them within their uncertainties. The effect on the differential cross section is typically $^{+1}_{-2}$%. Other uncertainties, related to the dependence on the average number of pile-up interactions of the efficiencies of the photon identification and isolation transverse energy requirements and to the observed data–MC shift in the photon isolation transverse energy distributions, are found to be negligible. A closure test has been performed by unfolding the differential spectra of di-photon events selected in the PYTHIA signal sample with the bin-by-bin coefficients determined using the SHERPA sample: non-closure effects of at most 2% have been found and included in the final systematic uncertainty. More sophisticated unfolding methods [149, 171] give results that are consistent with the nominal ones.

The integrated cross section is similarly measured by dividing the global $\gamma\gamma$ yield by the product of the average event selection efficiency, trigger efficiency and integrated luminosity. The selection efficiency, estimated to be $49.6^{+1.9}_{-1.7}$%, is computed from simulated di-photon events, reweighting the spectrum of one of the four di-photon variables under study in order to match the differential background-subtracted di-photon spectrum observed in data. The dominant contributions to the efficiency uncertainty originate from the photon identification efficiency uncertainty (±1.2%), the energy scale uncertainty ($^{+1.2}_{-0.5}$%), and the choice of the MC generator and the detector simulation (±0.9%). Negligible uncertainties are found to arise from the energy resolution, the isolation requirement (evaluated by shifting the isolation variable by the observed data–MC difference) and from the different pile-up dependence of the efficiency in data and MC simulation. The integrated cross section is $44.0^{+3.2}_{-4.2}$ pb, where the uncertainty is dominated by the systematic component stemming from the uncertainties on the event selection efficiency and the jet background subtraction.

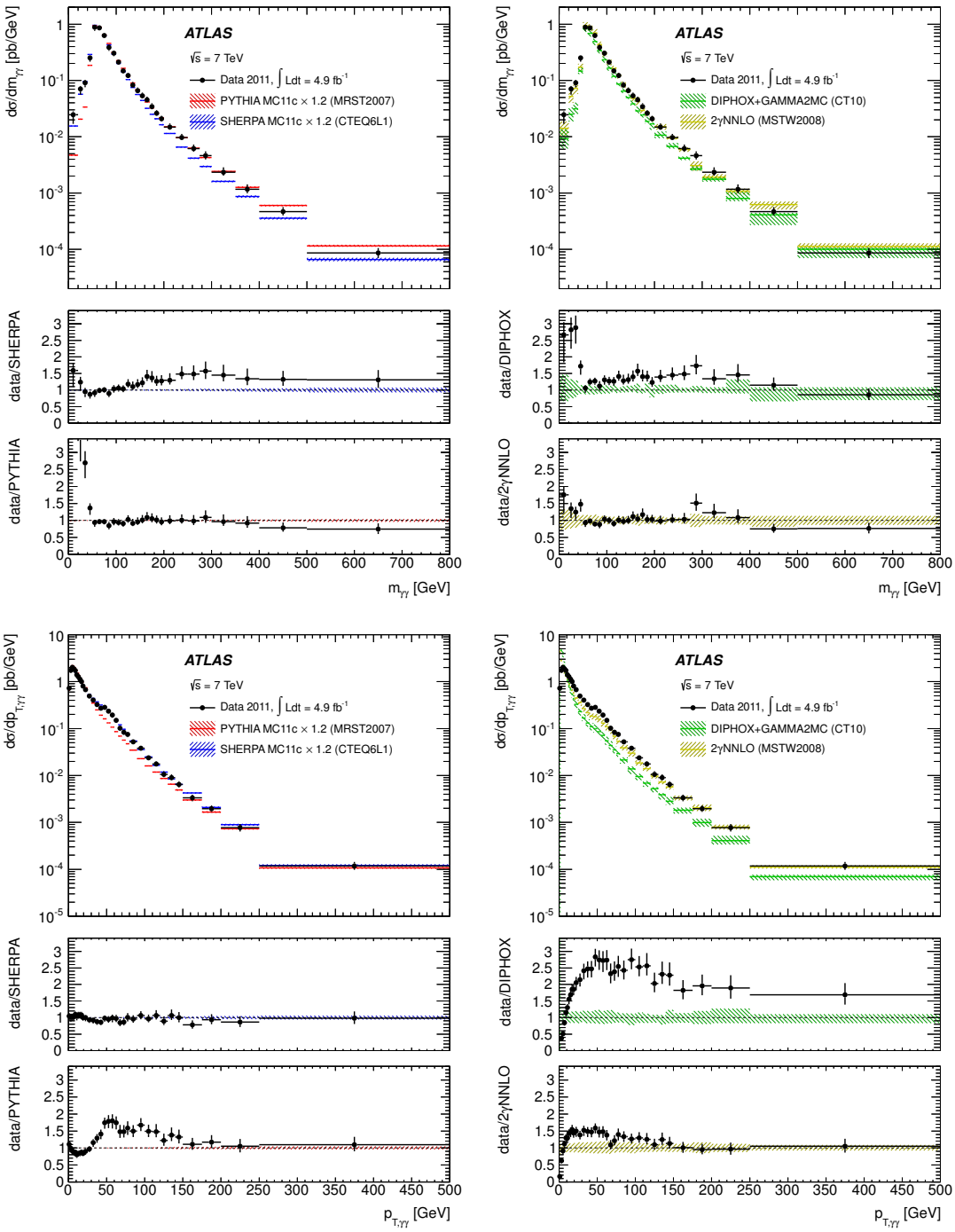

Figure 3.21: Comparison between the experimental differential di-photon cross sections and the predictions obtained with the parton-shower LO generators Sherpa and Pythia (left) or the parton-level generators, Diphox+gamma2mc (NLO) and 2γNNLO (NNLO), corrected for non-perturbative effects (right), as a function of $m_{\gamma\gamma}$ (top) or $p_{\mathrm{T},\gamma\gamma}$ (bottom).

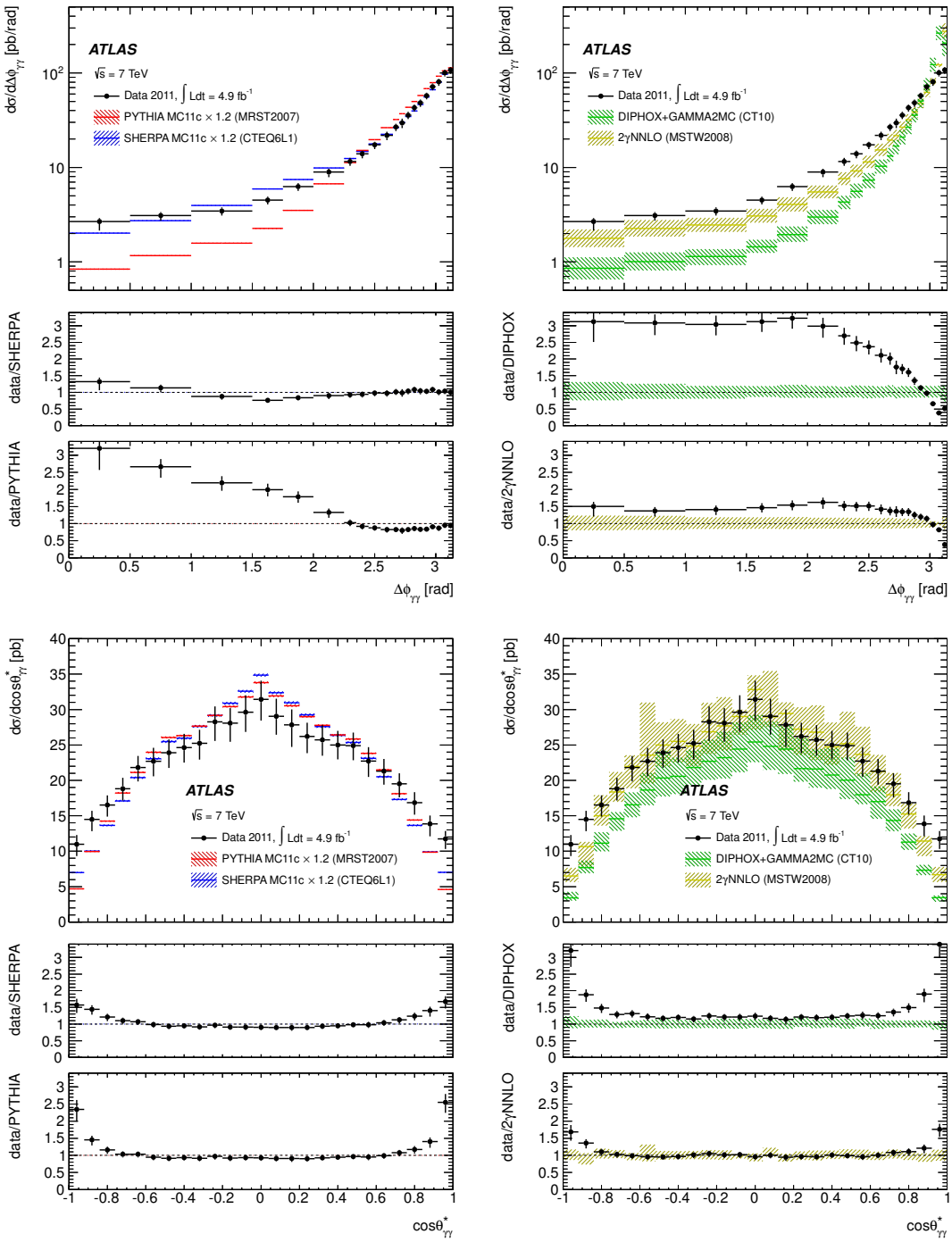

Figure 3.22: Comparison between the experimental differential di-photon cross sections and the predictions obtained with the parton-shower LO generators SHERPA and PYTHIA (left) or the parton-level generators, DIPHOX+GAMMA2MC (NLO) and 2γNNLO (NNLO), corrected for non-perturbative effects (right), as a function of $\Delta\phi_{\gamma\gamma}$ (top) or $\cos\theta^*_{\gamma\gamma}$ (bottom).

The total cross section estimated by Pythia and Sherpa with the ATLAS simulation settings is 36 pb, and underestimates the measured cross section by 20%. This is not surprising, as both Pythia and Sherpa lack NLO (and higher-order) contributions. The Diphox+gamma2mc total cross section is 39^{+7}_{-6} pb, about 10% lower than the measured one but in agreement with it within total uncertainties, and the 2γNNLO total cross section is 44^{+6}_{-5} pb, in good agreement with the data. The uncertainties on the theoretical predictions are dominated by the choice of the nominal scales. In Figs. 3.21 and 3.22, the differential cross sections measured in data and the corresponding predictions are presented. In order to compare the shapes of the Pythia and Sherpa differential distributions to the data, their cross sections are rescaled by a factor 1.2 to match the total cross section measured in data. The main features that can be noticed are the following:

- the Pythia cross section significantly underestimates the data for low values of $\Delta\phi_{\gamma\gamma}$, due to the missing higher-order contributions. On the other hand the parton shower is able to regularize the $\Delta\phi_{\gamma\gamma}$ cross section behaviour for $\Delta\phi_{\gamma\gamma}$ near π and at low $p_{\mathrm{T},\gamma\gamma}$.
- the shoulder expected (and observed) in the $p_{\mathrm{T},\gamma\gamma}$ cross section around the sum of the E_{T} thresholds of the two photons [172] is almost absent in Pythia, while Sherpa correctly reproduces the data in this region, as a consequence of the inclusion of the real emission NLO amplitudes combined with a different parton-shower model. Overall, Sherpa reproduces the data rather well, except at large $m_{\gamma\gamma}$ and large $|\cos\theta^*_{\gamma\gamma}|$.
- in the $\Delta\phi_{\gamma\gamma} \simeq \pi$, low $p_{\mathrm{T},\gamma\gamma}$ region, Diphox+gamma2mc fails to match the data as initial-state soft gluon radiation is divergent in fixed-order calculations without soft gluon resummation. Everywhere else Diphox+gamma2mc clearly underestimates the data due to the missing NNLO contributions.
- 2γNNLO, which includes NNLO corrections, is very close to the data within the uncertainties. However, the excess at $\Delta\phi_{\gamma\gamma} \simeq \pi$ and low $p_{\mathrm{T},\gamma\gamma}$ is still present, as expected for a fixed-order calculation. Since the fragmentation component is not calculated in 2γNNLO, the data is slightly underestimated by 2γNNLO in the regions where this component is larger: at low $\Delta\phi_{\gamma\gamma}$, low mass, intermediate $p_{\mathrm{T},\gamma\gamma}$ (between 20 GeV and 150 GeV) and large $|\cos\theta^*_{\gamma\gamma}|$.

As a consequence of the overall good agreement (up to a global scale factor) observed between Sherpa and the data in both the photon-jet and di-photon cross section measurements, the latest $H \to \gamma\gamma$ studies have used Sherpa instead of Pythia as the default background event generator.

Chapter 4

Search and discovery of the Higgs boson in decays to final states with photons

The search for the Higgs boson, culminated on July 4th 2012 with the discovery of a particle of mass around 125 GeV and properties (production rate, branching ratios, spin and parity, ...) consistent with those predicted by the Standard Model [27, 173], is one of the pillars of the LHC physics program. In this chapter, after briefly reviewing the expected production cross section and branching ratios of a SM Higgs boson at the LHC (Sec. 4.1) and the statistical methods used for the analysis of the data (Sec. 4.2), I will describe the Higgs boson searches ($H \to \gamma\gamma$, $H \to Z\gamma$) to which I have significantly contributed in the past years and the main results.

The discovery of the Higgs boson has been the successful result of a huge amount of collective work, performed by dedicated people in a well-coordinated way: it would have not been possible (to name a few key ingredients) without the tremendous performance achieved by the accelerator and the detector, the smooth data-taking conditions and the good quality of the data, the enormous computing power made available by the GRID infrastructure and the analysis tools developed within the collaboration, and of course the passion and ingenuity of hundreds of physicists working together on the analysis of the collected data, developing new ideas and cross-checking each others' results. As an example, the ATLAS $H \to \gamma\gamma$ working group was composed by around 85 physicists from 27 institutions at the time of the discovery; the LPNHE participation included several staff members as well as young Ph.D. students or post-docs.

My activities inside the ATLAS $H \to \gamma\gamma$ group, and more broadly on photon performance and physics aimed at the $H \to \gamma\gamma$ search, started in 2009 and have continued until nowadays. During these years, I worked on: the optimization of the photon identification criteria (Sec. 2.1) to maximize the $H \to \gamma\gamma$ signal efficiency for a given jet background rejection; $H \to \gamma\gamma$ sensitivity studies, described later in Sec. 4.3.1 (in particular, estimating the photon efficiency and jet rejection and the expected background from the simulation), before the start of the data-taking (2009 and early 2010); the measurement of various background cross sections (chapter 3) to validate the ATLAS simulation of prompt-photon events, to extract constraints on the gluon PDF (which is one of the largest sources of uncertainty on the theoretical Higgs boson cross section), and to develop data-driven techniques to estimate the separate $H \to \gamma\gamma$ background contributions (2010-2012); the early data analysis and first sensitivity studies, based on 38 pb^{-1} of 2010 data, focusing in particular on the determination of the background composition and comparison to the expectations, as described later in Sec. 4.3.2 (late 2010, early 2011). In spring 2011 I started to work on the photon identification efficiency measurements described in Sec. 2.2, in order to reduce the corresponding systematic uncertainty on the $H \to \gamma\gamma$ cross section measurement (or upper limit), and to develop for the LPNHE group, with two Ph.D. students, Heberth Torres and Olivier Davignon, a common $H \to \gamma\gamma$ analysis framework to perform event selection, Monte Carlo signal characterization, data-driven background composition measurements, and – thanks to the effort of Marine Kuna – the statistical treatment of the results and their interpretation in terms of upper

limits on the Higgs boson cross section normalized to the SM expectation. This allowed the LPNHE group to test frequently, during the second half of 2011 and the first half of 2012, different analysis strategies, evaluating the impact on the expected upper limit of changes in the selection, the photon identification and isolation criteria, or alternative choices of the event classification in categories; the quick feedback provided by our group and others helped the ATLAS $H \to \gamma\gamma$ working group to take decisions on the details of the analysis. In late 2011 and early 2012 I also collaborated with Heberth Torres and other LPNHE colleagues on studies of the $H \to \gamma\gamma$ background modeling and on how to reduce the bias induced from the assumed model on the estimated Higgs signal yield. The ATLAS $H \to \gamma\gamma$ analysis presented on July 4th 2012 and the corresponding results are summarized in Sec. 4.3.3. After the summer of 2012, once a Higgs boson had been discovered and the program of characterizing in detail this new particle, its properties and as many as possible final states was just starting, I began working on the search of the $H \to Z\gamma$ decays, leading with Rosy Nikolaidou the analysis on the full 7 and 8 TeV dataset, which is described, together with the main results, in Sec. 4.4.

4.1 Production cross sections and branching ratios for a Standard Model Higgs boson

The SM Higgs boson can be produced in pp collisions through five different processes involving the partons of the interacting protons, as illustrated in Fig. 4.1:

- gluon fusion, $gg \to H$
- vector-boson fusion (VBF), $q_1 q_2 \to q_3 q_4 H$
- associated production with a vector boson, $q_1 q_2 \to VH$, $V = W, Z$
- associated production with a $t\bar{t}$ pair, $gg \to t\bar{t}H$

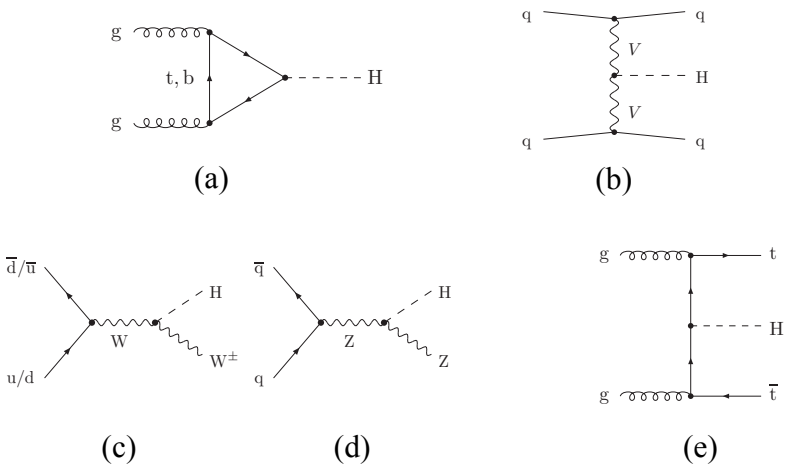

Figure 4.1: Main Feynman diagrams for the different Higgs boson production processes: gluon fusion (a), vector-boson fusion (b), associated production with a W (c) or Z (d) vector boson, and associated production with a $t\bar{t}$ pair (e).

Once produced, the Higgs boson can decay to a significantly larger number of final states, including final states containing photons like $H \to \gamma\gamma$ and $H \to Z\gamma$. As the Higgs boson does not couple directly to gluons nor to photons, the gluon fusion process and the decays to the $\gamma\gamma$ and $Z\gamma$ final states proceed only through loop diagrams, as illustrated in Figs. 4.1a and 4.2. The total amplitudes are dominated by the t-mediated (for $gg \to H$) or the W-mediated (for $H \to \gamma\gamma$ and $H \to Z\gamma$) contributions.

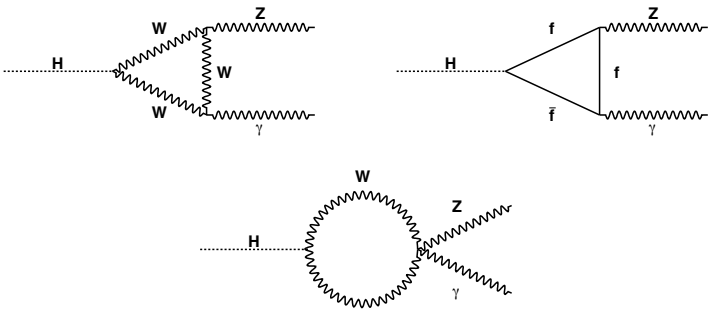

Figure 4.2: Leading Feynman diagrams for the $H \to Z\gamma$ decay in the Standard Model. In the case of the fermion loop, top quarks dominate. For $H \to \gamma\gamma$, the diagrams are similar, with the replacement $Z \to \gamma$.

For a given Higgs boson mass, knowing the parton momentum distribution functions inside the proton, the theoretical Higgs boson production cross sections, branching ratios and total width can be computed [174, 175, 176]. A compilation of the expected values, together with their uncertainties, is given in Refs. [42, 43]. The Higgs boson production cross sections are computed up to next-to-next-to-leading order (NNLO) [177, 178, 179, 180, 181, 182] in α_s for the gluon fusion process. The next-to-leading order (NLO) EW corrections are applied [183, 184]. These results are compiled in Refs. [185, 186, 187] assuming factorization between QCD and EW corrections. The cross sections for the VBF process are calculated with full NLO QCD and EW corrections [188, 189, 190], and approximate NNLO QCD corrections are applied [191]. The W/ZH processes are calculated at NLO [192] and at NNLO [193], and NLO EW radiative corrections [194] are applied. The full NLO QCD corrections for $t\bar{t}H$ are calculated [195, 196, 197, 198].

The Higgs boson production cross section as a function of the Higgs boson mass at CM energies of 7, 8 and 14 TeV, and the Higgs boson branching ratios for different decay channels including those under study, are presented in Fig. 4.3. Raising the pp center-of-mass energy from 7 to 8 TeV increases the production cross section by almost 30%, while going from 8 to 14 TeV the cross section increases by ≈ 2.6. At $\sqrt{s} = 8$ TeV the production cross section for a Higgs boson with a mass of 125 GeV is $\sigma_H = 22.3$ pb; the contributions from gluon fusion, VBF, VH and $t\bar{t}H$ are respectively 87%, 7%, 5% and 1%. At the same mass, the relative uncertainties on the gluon fusion and $t\bar{t}H$ cross sections are around 15% (about 8% each from the scales and the PDF uncertainties) while those on the VBF and VH cross sections are around 5% (1% from the scales and 4% from the PDFs). The branching ratios for $H \to \gamma\gamma$ and $H \to Z\gamma$ are 2.28×10^{-3} and 1.54×10^{-3}, with relative uncertainties of 5% and 9%, respectively, due to the uncertainties on the QCD energy scales and on the values of the quark masses and of the strong coupling constant.

4.2 Statistical methods

The significance of the observed signal or the upper limit on its effective cross section are set using the profile likelihood technique, as described in detail in Ref. [199]. The results are expressed in terms of a "signal-strength" parameter μ, defined as the ratio

$$\mu = \frac{\sigma \times BR}{(\sigma \times BR)_{\mathrm{SM}}} \tag{4.1}$$

between the measured and expected effective cross section (including the branching ratio to the final state of interest), and equal to the ratio of the measured number of signal events to the value expected in the Standard Model, $\mu = \frac{N_{\mathrm{signal}}}{N^{\mathrm{SM}}_{\mathrm{signal}}}$. To constrain the value of μ, the distribution of an observable x that discriminates between signal and background events is fitted with a global

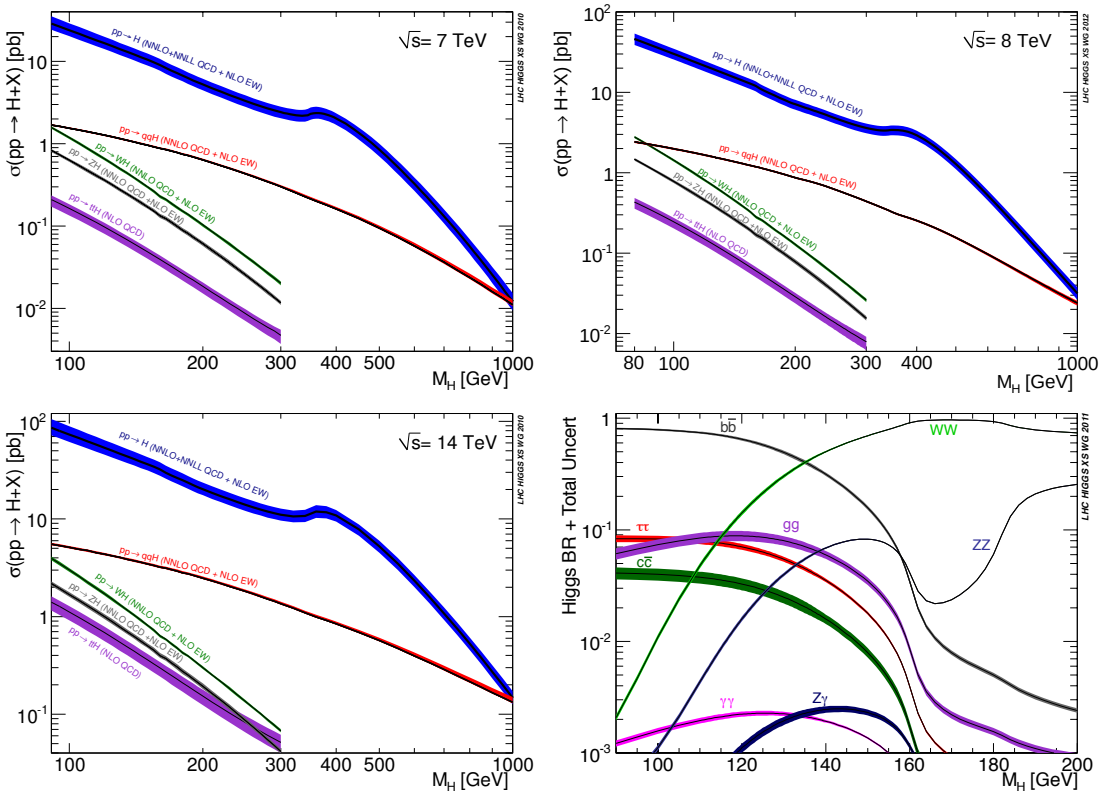

Figure 4.3: SM Higgs boson production cross section for the different production mechanisms vs Higgs boson mass at $\sqrt{s} = 7$ TeV (top-left), 8 TeV (top-right) and 14 TeV (bottom-left), and SM Higgs boson branching ratios vs Higgs boson mass, for various final states (bottom-right) [42, 43].

likelihood which is the product of per-event likelihoods of the form

$$\mathcal{L}(\mu, \boldsymbol{\theta}|x) = \frac{\mu N_S^{\mathrm{SM}}(\boldsymbol{\theta})}{\mu N_S^{\mathrm{SM}}(\boldsymbol{\theta}) + N_B(\boldsymbol{\theta})} f_S(x|\boldsymbol{\theta}) + \frac{N_B(\boldsymbol{\theta})}{\mu N_S^{\mathrm{SM}}(\boldsymbol{\theta}) + N_B(\boldsymbol{\theta})} f_B(x|\boldsymbol{\theta}), \tag{4.2}$$

where $\boldsymbol{\theta}$ represent the sets of additional (*nuisance*) parameters on which the signal and background distributions and yields may depend on. The global likelihood for the dataset under study is thus:

$$L(\mu, \boldsymbol{\theta}|\mathbf{x}) = \left(e^{-N'} N'^N\right) \prod_{k=1}^{N} \mathcal{L}(\mu, \boldsymbol{\theta}|x_k) \tag{4.3}$$

where $\mathbf{x} = \{x_1, .., x_N\}$ is the set of the measured values of the variable x in the N selected events in data, and a Poisson probability factor for the observed number of events has also been included. In case the selected events are classified in a number n_{cat} of orthogonal categories c, the full likelihood is the product of n_{cat} likelihoods L_c of the previous form, one for each category:

$$L\left(\mu, \boldsymbol{\theta} = \bigcup_{c=1}^{n_{cat}} \boldsymbol{\theta_c} \middle| \mathbf{x} = \bigcup_{c=1}^{n_{cat}} \mathbf{x_c}\right) = \prod_{c=1}^{n_{cat}} L_c(\mu, \boldsymbol{\theta_c}|\mathbf{x_c}) \tag{4.4}$$

and the per-event likelihood for category c is

$$\mathcal{L}_c(x|\mu, \boldsymbol{\theta_c}) = \frac{\mu N_{S,c}^{\mathrm{SM}}(\boldsymbol{\theta_c})}{\mu N_{S,c}^{\mathrm{SM}}(\boldsymbol{\theta_c}) + N_{B,c}(\boldsymbol{\theta_c})} f_{S,c}(x|\boldsymbol{\theta_c}) + \frac{N_{B,c}(\boldsymbol{\theta_c})}{\mu N_{S,c}^{\mathrm{SM}}(\boldsymbol{\theta_c}) + N_{B,c}(\boldsymbol{\theta_c})} f_{B,c}(x|\boldsymbol{\theta_c}) \tag{4.5}$$

Finally, auxiliary measurements that help constrain some of the nuisance parameters may be available; in the typical case, one has a best guess a_p of the parameter θ_p and some relative uncertainty σ_p on that estimate. The likelihood is thus multiplied by a "constraint term" $f_p(a_p|\theta_p, \sigma_p)$ for the "global" observable a_p. If the subset of parameters with constraint terms is denoted S and $\mathbf{a}$ are the corresponding global observables, the full likelihood becomes, including the constraint terms explicitly:

$$L\left(\mu, \boldsymbol{\theta}|\mathbf{x}, \mathbf{a}\right) = \prod_{c=1}^{n_{cat}} L_c(\mu, \boldsymbol{\theta_c}|\mathbf{x_c}) \prod_{p \in S} f_p(a_p|\theta_p, \sigma_p) \tag{4.6}$$

The profile likelihood ratio for the parameter of interest μ is then defined as the ratio

$$\Lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{4.7}$$

between the "conditional" maximum value of the likelihood L for a fixed value of μ, obtained when the nuisance parameters take the values $\hat{\hat{\boldsymbol{\theta}}}(\mu)$, and the maximum value of the likelihood, obtained for $\mu = \hat{\mu}, \boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. If the parameter of interest μ represents a non-negative quantity, like in the case of a signal event rate, the profiled likelihood ratio is modified as

$$\Lambda(\mu) = \begin{cases} \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} & \hat{\mu} \geq 0 \\ \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(0, \hat{\hat{\boldsymbol{\theta}}}(0))} & \hat{\mu} < 0 \end{cases} \tag{4.8}$$

In order to quantify the significance of a possible observation, a hypothesis test is performed to evaluate the compatibility between the data and the background-only (B-only) hypothesis ($\mu = 0$) and to discriminate the $\mu = 0$ and $\mu > 0$ hypotheses. The $\tilde{q}_0$ test statistic is used:

$$\tilde{q}_0 = \begin{cases} 0 & \hat{\mu} < 0 \\ -2 \ln \Lambda(0) & \hat{\mu} \geq 0 \end{cases} \tag{4.9}$$

From the previous definitions, $\tilde{q}_0$ is non negative; it is zero if the data favors the background-only hypothesis ($\hat{\mu} \leq 0$) and increases the further the fitted $\hat{\mu}$ is from zero. The p-value of the null hypothesis is

$$p_0 = \int_{\tilde{q}_{0,\mathrm{obs}}}^{+\infty} f(\tilde{q}_0|0, \hat{\hat{\boldsymbol{\theta}}}(0)) d\tilde{q}_0, \tag{4.10}$$

where f is the distribution of the test statistics. It quantifies the probability for a dataset generated in the B-only hypothesis to have a $\tilde{q}_0$ larger than the observed one, *i.e.* the probability to be in the same or worse agreement with the null hypothesis than the data. In the previous formula, we made explicit the fact that the true distribution $f(\tilde{q}_0|\mu = 0)$ is to some extent unknown, as the values of the nuisance parameters are unknown, and one has thus to make some assumptions on them; in principle one may compute the p_0 for any value of $\boldsymbol{\theta}$ and take the maximum p_0, but this would be practically unfeasible. However, asymptotically – when N is large – the distribution of the profile likelihood ratio is independent of the values of the nuisance parameters, and so is the p_0; in our case, we estimate it by fixing the values of the nuisance parameters to those that maximise the likelihood for $\mu = 0$. The closed-form asymptotic formulae for $f(\Lambda(\mu)|\mu)$ and $f(\Lambda(\mu)|\mu')$ are known [200] and can be used to compute the p_0 in the asymptotic regime; alternatively, the p_0 can be computed by sampling the distribution of $\tilde{q}_0$ in the B-only hypothesis using pseudo-experiments. The p_0 is usually converted into a *significance* computed as the quantile or number of σ of a unit Gaussian,

$$Z = \Phi^{-1}(1 - p_0), \tag{4.11}$$

where Φ^{-1} is the inverse of the cumulative distribution for a unit Gaussian.

Upper limits on the signal strength are set using a modified frequentist (CL_s) [201] method, using a test statistic $\tilde{q}_\mu$ to differentiate between the μ and $\mu' > \mu$ hypotheses:

$$\tilde{q}_\mu = \begin{cases} 0 & \mu < \hat{\mu} \\ -2\ln\Lambda(\mu) & \mu \geq \hat{\mu} \end{cases} = \begin{cases} 0 & 0 \leq \mu < \hat{\mu} \\ -2\ln\frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(0,\hat{\hat{\boldsymbol{\theta}}}(0))} & \hat{\mu} < 0 \leq \mu \\ -2\ln\frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu},\hat{\boldsymbol{\theta}})} & 0 \leq \hat{\mu} \leq \mu \end{cases} \tag{4.12}$$

This test statistic, defined for $\mu \geq 0$, is thus zero for μ below the best-fit value $\hat{\mu}$, and increases as the difference $\mu - \hat{\mu}$ becomes larger. The observed p-value of μ in the $\mu S + B$ hypothesis (CL_{s+b}) is defined as:

$$p_\mu = \int_{\tilde{q}_{\mu,\text{obs}}}^{+\infty} f(\tilde{q}_\mu|\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))d\tilde{q}_\mu \tag{4.13}$$

and represents the probability that a dataset generated according to the $\mu S + B$ hypothesis will have a value of $\tilde{q}_\mu$ greater than that observed in data. The CL_s p-value, used to set upper limits, is defined as

$$CL_s(\mu) = \frac{p_\mu}{1 - p_b} \tag{4.14}$$

where $1 - p_b$ is the p-value (CL_b) derived from the same test statistic under the background-only hypothesis,

$$1 - p_b = \int_{\tilde{q}_{\mu,\text{obs}}}^{+\infty} f(\tilde{q}_\mu|0, \hat{\hat{\boldsymbol{\theta}}}(0))d\tilde{q}_\mu \tag{4.15}$$

In practice, CL_{s+b} quantifies the probability for the sum of signal and background ($\mu S + B$) to fluctuate downwards to yield an observed $\mu = \hat{\mu}$; CL_s normalizes CL_{s+b} by the probability for the background to fluctuate (upwards) to yield $\mu = \hat{\mu}$. CL_s thus allows the experimenter to extract a statement, with an approximate confidence, on the signal hypothesis, unlike CL_{s+b} which should be interpreted as a statement on the total of signal and background. The value of $CL_s(\mu)$ and of the corresponding exclusion can be obtained using either asymptotic formulae [200] or pseudo-data generation. Limits at 95% confidence level (C.L.) on the value of the signal strength μ are computed by scanning values of the μ hypothesis, computing the corresponding CL_s and identifying the value μ_{up} for which CL_s equals 0.05.

For sensitivity studies, the expected p_0 (assuming the Standard Model signal rate) and expected μ_{up} (assuming the background-only hypothesis) are also computed. They correspond to the median values of the distributions $f(p_0|\mu = 1, \boldsymbol{\theta})$ and $f(\mu_{up}|\mu = 0, \boldsymbol{\theta})$, where the nuisance parameters are fixed to the profiled values based on the observed data. 1σ and 2σ uncertainty bands can also be computed from these distributions. These values can be evaluted either by using an ensemble of generated pseudo-experiments to obtain the full distribution f, or computing the observed p_0 and μ_{up} for a single representative "Asimov" dataset, *i.e.* a special generated sample such that when one uses it to evaluate the estimators for all parameters, one obtains the true parameter values [200].

4.3 Observation of a Higgs-like boson decaying to di-photons, $H \to \gamma\gamma$

While a light ($m_H \approx 125$ GeV) Standard Model Higgs boson can decay to several final states, many of them are either overwhelmed by a huge QCD background (gg, $c\bar{c}$, $b\bar{b}$) or are characterized by a poor invariant mass resolution, either due to the limited $E_{\mathrm{T}}^{\mathrm{miss}}$ (WW, $\tau\tau$) or jet ($b\bar{b}$) energy resolution. With an estimated $S/B \approx 1.5$ and an invariant mass resolution below 2%, the $H \to ZZ^* \to 4\ell$ decay is the golden channel for the discovery of a light SM Higgs boson at the LHC. However, reconstructing the Z boson in light di-lepton final states ($BR(Z \to \ell\ell) = 6.7\%$) reduces the overall cross section to 2.6 fb, and only a few events are expected in about 20 fb^{-1} at 8 TeV when including selection efficiencies. For this reason, the di-photon decay is equally relevant, as it can achieve a similar invariant mass resolution, thanks to the excellent energy resolution of the electromagnetic calorimeter, but much larger yield, due to an effective cross section (including $BR(H \to \gamma\gamma) = 2.28 \times 10^{-3}$ at $m_H = 125$ GeV) of 51 fb at $\sqrt{s} = 8$ TeV, though the expected S/B – even with an excellent jet rejection – is only around 3%, mostly because of the irreducible SM di-photon background.

One of the goals of the ATLAS LAr calorimeter design was indeed to allow the detection of a significant $H \to \gamma\gamma$ signal in pp collisions at $\sqrt{s} = 14$ TeV. The first studies of the discovery potential of a SM Higgs boson in the mass range $80 \lesssim m_H \lesssim 140$ GeV, in the di-photon decay channel, with an "ATLAS-like" LAr electromagnetic calorimeter, date back to 1991 [45], before the proposed EAGLE (Experiment for Accurate Gamma, Lepton and Energy Measurements) and ASCOT (Apparatus with Super Conducting Toroids) collaborations merged into the ATLAS collaboration, in 1992. These studies were based on the following assumptions:

- $\sqrt{s} = 16$ TeV (initial proposed LHC CM energy)
- $\int Ldt = 100$ fb^{-1} (1 year of LHC at design luminosity)
- a LAr sampling electromagnetic calorimeter with the following characteristics:
  - a barrel ($|\eta| < 1.42$) and an end-cap ($|\eta| < 2.5$) region, separated by a crack ($\Delta\eta = 0.2$).
  - an energy resolution $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 1\% \oplus \frac{0.3}{E}$ (E in GeV).
  - an angular resolution $\sigma_\theta = \frac{100 \text{ mrad}}{\sqrt{E}}$. For a resolution like this one or better (or, equivalently, a resolution on the longitudinal coordinate of the photon production vertex, σ_z, better than 1.5 cm), the di-photon invariant mass resolution is dominated by the energy resolution. With the energy resolution given above, the invariant mass resolution for $m_H = 120$ GeV is $\sigma_m/m = 1.21\%$ ($\sigma_m = 1.34$ GeV).
  - a finely segmented (≈ 3 mm) presampler with a depth $X/X_0 \approx 3$ to discriminate single photons from isolated π^0s.

A simple inclusive search was foreseen, requiring the reconstruction of two photons with transverse energies above 40 GeV and 25 GeV, respectively (to reduce background and trigger rate), and a small transverse momentum imbalance, $\frac{p_{\mathrm{T},1}}{p_{\mathrm{T},1}+p_{\mathrm{T},2}} < 0.7$ (to reduce background from fragmentation-induced SM di-photon events). The kinematic acceptance for $m_H = 120$ GeV was 45%. The significance was estimated by computing the ratio $S/\sqrt{B}$ from the expected number of signal (S) and background (B) events in an optimized di-photon mass bin of width $\pm 1.25\sigma_m$ ($\varepsilon \approx 79\%$). The inputs used for the calculation were the following:

- the signal production cross section was computed for the two leading processes, gluon fusion and vector-boson fusion.
- as the backgrounds cross sections were known only at LO in QCD, both signal and background cross sections were consistently computed at LO (though the NLO QCD corrections for the signal were available). For the signal, this gave $\sigma_H \times BR_{\gamma\gamma} = 71$ fb (before acceptance).
- the reconstruction and identification efficiency, for photons in the acceptance, was assumed to be $\varepsilon_\gamma = 80\%$.

- the jet rejection, combining information from leakage in the hadronic compartment, isolation, shower shape in the calorimeter and shower transverse profile in the presampler, was estimated with a fast simulation of the EAGLE calorimeter to be around $R_j = 10000$. As the ratios between the background cross sections in the acceptance are of the order $\frac{\sigma_{\gamma j}}{\sigma_{\gamma\gamma}} \approx 10^3$ and $\frac{\sigma_{jj}}{\sigma_{\gamma\gamma}} \approx 6 \times 10^6$, the final background was expected to be dominated by SM di-photon events, with $\frac{N_{\gamma j}}{N_{\gamma\gamma}} \approx \frac{10^3}{R_j \varepsilon_\gamma} = 12.5\%$ and $\frac{N_{jj}}{N_{\gamma\gamma}} \approx \frac{6\times10^6}{R_j^2 \varepsilon_\gamma^2} = 9\%$. The background yield B was thus estimated neglecting completely the γ-jet and di-jet components.

The calculation yielded $S = 1640$, $B = 2.2 \times 10^4$ and $S/\sqrt{B} = 11.1$ in the optimized mass bin.

These studies were later updated for the ATLAS Technical Proposal in 1994 [202] and during the final stages of the detector installation in the ATLAS cavern and before the expected start of the operations in 2008 [203], assuming that the LHC would run at a center-of-mass energy of 14 GeV, delivering 10 fb^{-1} during the first year in the initial, low-luminosity (10^{33} cm^{-2} s^{-1}) phase, and then 100 fb^{-1} per year at the design luminosity. In particular, Ref. [203] used the NLO cross sections for the signal (including all production decay modes) and for the three main background processes ($\gamma\gamma$, γj, jj), and a more realistic (though fast) simulation of the detector to better assess the expected performance of ATLAS in terms of photon efficiency (estimated to be $\varepsilon_\gamma \approx 81\%$) and jet rejection ($R_j \approx 8000$) using photon identification and isolation requirements. Assuming a 1-year integrated luminosity of 10 fb^{-1} and a Higgs boson mass of 120 GeV, this study anticipated a 2.6σ significance (without taking into account the trial factor or "look-elsewhere" effect) with a cut-and-count analysis and a 3.5σ significance using a more sophisticated statistical treatment of the data based on a likelihood fit to the di-photon invariant mass distribution and the classification of the data in orthogonal categories with different S/B.

4.3.1 MC-based feasibility studies at $\sqrt{s} = 10$ TeV and 7 TeV

The delay of the building of the LHC and of the start of the *pp* collisions, the progressive lowering of the initial expected *pp* center-of-mass energy, and the excellent performance of the CDF and D0 experiments to collect and analyse the $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV delivered by the TeVatron accelerator to search for a light Higgs boson decaying to several final states, particularly to a $b\bar{b}$ pair, prompted ATLAS to update frequently its Higgs boson sensitivity studies, in order to provide projections of the expected exclusion potential for different luminosity and CM energy scenarios. These inputs, together with those from CMS, were then used by the CERN management, upon discussion with the LHC accelerator and experiment representatives, to decide the LHC schedule for the years 2010-2012.

The initial LHC running plan for 2010, proposed at the Chamonix workshop in February 2009, was to start its physics run as soon as the repair activity after the accident of September 2008 would be over and to collect around 200 pb^{-1} of *pp* collisions at $\sqrt{s} = 10$ TeV, before the shutdown foreseen at the end of 2010 [204]. In 2009, therefore, within an ATLAS-wide readiness effort aiming to prepare the analysis of real data to be collected in 2010, a few ATLAS groups worked on updating the previous sensitivity studies, using the latest fully simulated signal and background Monte Carlo samples, state-of-the-art theoretical calculations, and realistic estimates of systematic uncertainties to evaluate the exclusion potential for a SM Higgs boson decaying to two photons with such a luminosity. In particular, the effort was started and lead by four French groups (LAPP/LAL/LPNHE/LPSC), collaborating to the French ANR programme "Higgsnet", and joined by a few other groups.

Meanwhile, during the repair and consolidation work of 2009, the LHC physicists arrived at the conclusion that for 2010, operation at an energy of 5 TeV per beam would have been risky, while 3.5 TeV per beam would have been a safe option. At the Chamonix workshop in January 2010 it was thus decided to operate the LHC at $\sqrt{s} = 7$ TeV during 2010 with the goal of collecting 1 fb^{-1} of *pp* collisions, before a rather long shutdown to consolidate the whole machine for an energy of 7 TeV per beam [205]. The Higgs boson sensitivity studies were thus updated to reflect the lower center-of-mass region.

These pre-data-taking $H \to \gamma\gamma$ sensitivity studies are documented in an internal ATLAS note [206] and are part of a public document covering a wider spectrum of Higgs boson final states [19]. The results are based on a full simulation of the signal (for $m_H = 120$, 125 and 130

GeV) and of the main backgrounds (di-photon, photon-jet, di-jet and Drell-Yan). The signal samples are generated at both $\sqrt{s} = 10$ and 7 TeV, while the background samples are generated only at $\sqrt{s} = 10$ TeV, which was the foreseen LHC center-of-mass energy when this study started; the results are then extrapolated to $\sqrt{s} = 7$ TeV using the ratios of the background production cross sections at the two CM energies. For consistency, both signal and background samples, which are produced using LO parton shower MC generators, are normalised to the NLO calculations of the corresponding cross sections. A low instantaneous luminosity is assumed and no pile-up is included in the simulation.

The search is based on a deliberately simple and robust analysis, using only the reconstructed di-photon mass as a discriminating variable, in order to minimise the impact of systematic uncertainties. Events are required to pass a loose di-photon trigger with symmetric E_{T} thresholds at 20 GeV, and to contain two photons within the fiducial region of the electromagnetic calorimeter for photon identification ($|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$) and with transverse momenta greater than 40 and 25 GeV, respectively. The photons must pass tight identification requirements and a track-isolation requirement, $\sum_{\mathrm{trk}} p_{\mathrm{T}}^{\mathrm{trk}} < 4$ GeV, where the tracks included in the sum must lie within $0.1 < \Delta R < 0.3$ from the photon, have $p_{\mathrm{T}} > 1$ GeV, at least one b-layer hit and 7 hits in the silicon detectors, and a transverse impact parameter smaller than 1 mm. The isolation requirement has a 99% efficiency on the signal and a background rejection factor around 1.5.

The di-photon invariant mass is reconstructed from the photon energies and directions; the latter are determined using the estimated impact points of the photons in the first layer of the ECAL, assuming as origin along the beam axis the point which maximises a likelihood that combines the measured z coordinate of the event primary vertex (the one with largest $\sum p_{\mathrm{T}}^2$ of the associated tracks) and the z coordinate of the di-photon vertex. This vertex is computed as the weighted average of the z coordinates of the intersections with the beam line of the photon directions (in the $R - z$ plane) obtained using calorimeter pointing (exploiting the measurements in the first and second layers of the ECAL) and the conversion vertex position in case of a conversions in the silicon detectors.

For $m_H = 120$ GeV the signal efficiency according to the Pythia simulation is around 46% at both 7 and 10 TeV and the expected number of signal events is 5 for a luminosity of 200 pb^{-1} at 10 TeV and 13 for a luminosity of 1 fb^{-1} at 7 TeV. The SM di-photon yield and invariant mass distribution are obtained from fully simulated Alpgen [207] (for the Born and bremsstrahlung contributions) and Pythia (for the box contribution) samples after reweighting the $p_{\mathrm{T}}^{\gamma\gamma}$ spectra and scaling the cross sections according to NLO predictions. For the reducible backgrounds, the small number of fully simulated events (few millions) compared to the large jet rejection of the photon identification requirements leads to large uncertainties on the invariant mass spectrum because of large statistical fluctuations in the few events that pass the full selection. To overcome this limitation, we developed an alternative method, in which we first extract parametrizations, as a function of the true object transverse momentum, of both the photon identification efficiency and the quark-jet and gluon-jet rejection factors R (Sec. 2.1.4), the latter being determined as the ratio between the true p_{T} spectra of jets associated to reconstructed photons and the true p_{T} spectra of all jets; then, using fully simulated samples, we extract a parametrization of the jet momentum fraction carried by the reconstructed fake photon, $p_{\mathrm{T}}^{\gamma}/p_{\mathrm{T}}^{\mathrm{jet}}$, as a function of the true $p_{\mathrm{T}}^{\mathrm{jet}}$, separately for quarks and gluons. Finally, we use these parametrizations to weight generated di-jet and photon-jet events, thus obtaining the final estimates of the reducible background. As shown in Fig. 4.4, for the photon-jet background a good agreement with the distribution from the full simulation is found, and a much smoother prediction of the $m_{\gamma\gamma}$ distribution is obtained; for the di-jet background, this method allows us to estimate the number of events expected in the range $100 < m_{\gamma\gamma} < 150$ GeV, which is not possible otherwise as no event of the di-jet simulated sample passes the full selection because of the small equivalent luminosity, 0.5 pb^{-1}.

An example of the expected invariant mass distribution with 1 fb^{-1} of data at $\sqrt{s} = 7$ TeV is shown in Fig. 4.5. About 8500 background events are expected in the range $100 < m_{\gamma\gamma} < 150$ GeV: 65% are due to SM di-photon, 30% to photon-jet, 4% to di-jet and 1% to Drell-Yan events. The background distribution can be described by a simple exponential function, whose slope is extracted through a maximum likelihood fit to the data. This function and the di-photon invariant mass distribution of signal events, parametrized with the sum of a Crystal Ball lineshape [208], describing the core resolution and a non-Gaussian tail towards lower mass values, and a Gaussian contribu-

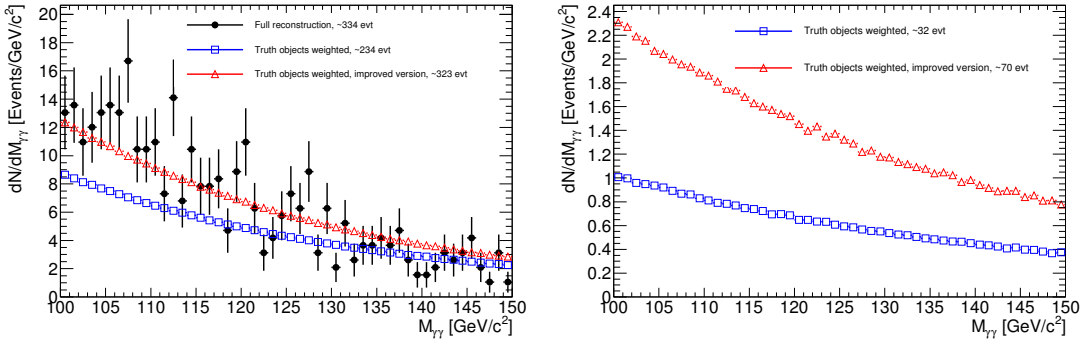

Figure 4.4: Di-photon candidate invariant mass spectra for $\gamma-$jet events (left) and di-jet events (right) obtained from fully reconstructed MC samples (dots with error bars) or from generator-level samples, using the efficiencies and rejection factors measured as a function of transverse momentum on fully reconstructed samples (red triangles) or using integrated efficiencies and rejection factors (blue squares, for comparison). The distributions are normalised to a luminosity of 200 pb^{-1} at $\sqrt{s} = 10$ TeV, using the LO cross sections provided by the MC generator.

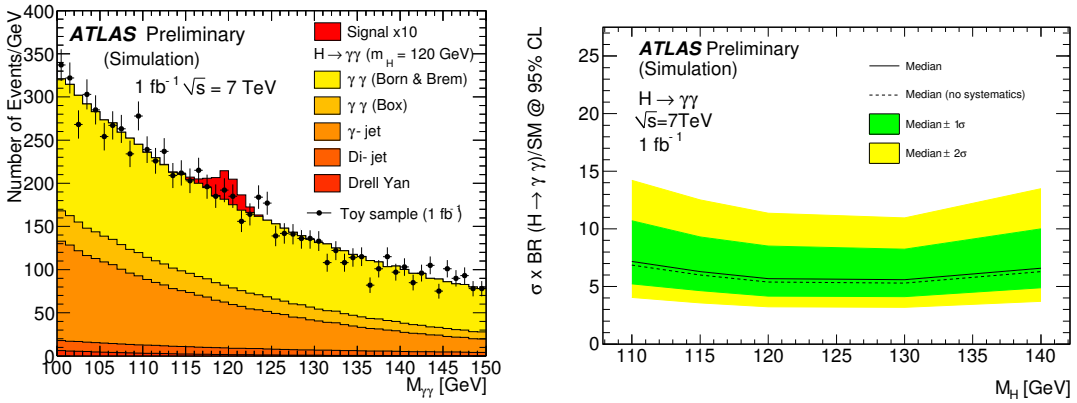

Figure 4.5: Left: expected invariant mass distribution of di-photon candidate events selected at $\sqrt{s} = 7$ TeV for an integrated luminosity of 1 fb^{-1}. The SM Higgs boson contribution is enhanced by a factor 10. Right: estimated median number of Standard Model signal cross-section excluded at 95% C.L. as a function of the Higgs mass for the same amount of data, using the CL_s technique. The green and yellow bands represent the range in which we expect the observed limit to lie (at 68% and 95% C.L. respectively), depending on the data.

tion describing the outliers, are then used to generate background-only and signal+background pseudodata, from which the distributions of the test statistics $\tilde{q}_\mu$ in the two cases and the expected 95% C.L. upper limit on the effective cross section are extracted. As an example, the expected CL_s upper limit on μ with 1 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV is shown in Fig. 4.5. With 200 pb^{-1} (1 fb^{-1}) at $\sqrt{s} = 10$ (7) TeV, the expected upper limit on μ is $\mu < 7.2$ (5.6) for $m_H = 120$ GeV, significantly better than the limits set by CDF (22.5) [209] and D0 (19.4) [210] at the time of this study, for the same Higgs mass and the same final state.

A few systematic uncertainties on the expected limit at 7 TeV have been investigated: the effects of a larger constant term of the photon resolution (1.1% instead of 0.7%), of a 1% photon efficiency uncertainty and of a 10% luminosity uncertainty have been estimated as 9%, 2% and 10%, respectively. An additional ±13% uncertainty on the expected upper limit is due to the uncertainties on the di-photon and photon-jet composition, which affect the final $m_{\gamma\gamma}$ background distribution used to generate the pseudo-data for the sensitivity studies.

In conclusion, including uncertainties in the extrapolations (and even rescaling the limits by ≈ 1.2 to account for a 20% uncertainty on the theoretical predictions), both initial running scenarios of the LHC would allow ATLAS to set upper limits on the effective cross section of a light Higgs decaying to di-photons competitive with those of the Tevatron experiments.

4.3.2 Early data analysis: measurement of the backgrounds to the $H \to \gamma\gamma$ search, reappraisal of its sensitivity and first upper limits on the production cross section.

In winter 2010, after the LHC had delivered its first ≈ 40 pb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV, the ATLAS $H \to \gamma\gamma$ working group exploited this data to:

- determine the background composition and compare it to the predicted one,
- perform a first search of the $H \to \gamma\gamma$ decay with the ATLAS data,
- extract the $m_{\gamma\gamma}$ background distribution for extrapolations of the analysis sensitivity at higher integrated luminosities.

The main features of the analysis and the results [20, 21] are briefly summarized here. The selection requirements and the analysis strategy are to a large extent similar to those outlined in the previous section: events are collected with a loose di-photon trigger with symmetric E_{T} thresholds at 15 GeV and are required to contain two photons with pseudorapidity $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$ and with transverse momenta greater than 40 and 25 GeV, respectively. The photons must pass tight identification requirements and a calorimeter isolation requirement, $E_{\mathrm{T}}^{\mathrm{iso}} < 3$ GeV in a cone of radius 0.4 around the photon, as used in the measurements of the single- and di-photon cross sections with the same data. The calorimeter isolation is used instead of the track isolation exploited in previous studies as it is found to be more effective in reducing the jet background and as pile-up robust (after the energy-density correction described in Sec. 1.6.5) as track isolation. In total, 99 events with a di-photon invariant mass between 100 and 150 GeV are selected, for a luminosity of 38.0 ± 1.3 pb^{-1}. The number of expected signal events for a SM Higgs boson with a mass of 120 GeV is $0.45^{+0.11}_{-0.10}$.

The number of di-photon, photon-jet and di-jet events are extracted with two methods, a double-sideband counting one and a two-dimensional isolation fit similar to that described in Sec. 3.7. The photon templates are determined from the $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution of electrons from W and Z decays, shifted to compensate for the differences between electrons and photons seen in simulation. The fake-photon isolation templates (for photon-jet events) are obtained by reversing a subset of the identification requirements, separately for leading and sub-leading photon candidates. The di-jet 2D template is obtained by reversing the identification requirements simultaneously for the two photon candidates. Before the fit to the data, a closure test is performed on a simulated sample containing a mixture of di-photon and photon-jet events corresponding to an integrated luminosity of about 110 pb^{-1}; no significant bias is observed comparing the fit results to the true yields. The results of the two methods agree with each other, and, as shown in Fig. 4.6, they agree with the expectations, within the large experimental (18%, 42%, 100%) and theoretical (27%, 50%, 100%) uncertainties for the di-photon, photon-jet and di-jet backgrounds, respectively. The

experimental errors are dominated by the statistical component.

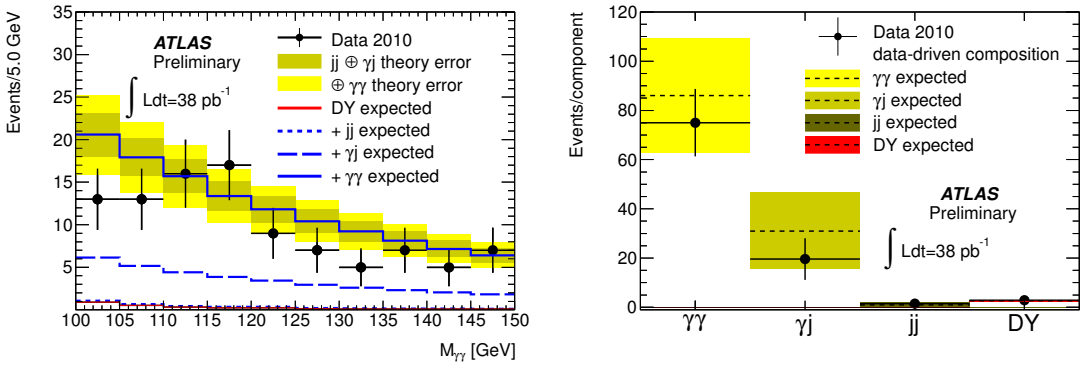

Figure 4.6: Left: di-photon invariant mass distribution of the 99 events selected in 2010 at $\sqrt{s} = 7$ TeV with an integrated luminosity of 38 pb^{-1}. The overlaid histograms represent cumulative Drell-Yan (red solid), di-jet (blue dotted), photon-jet (blue dashed) and diphoton (blue solid) components of the background, according to the predictions from theoretical models and simulation. The dark yellow band is the uncertainty for the reducible background components, and the yellow band is the total uncertainty on the reducible plus irreducible backgrounds. Right: comparison between the estimated background yields on data (black dots) with the corresponding predictions. For the Drell-Yan component, the number of events is compared with the expected number of events predicted from the full simulation.

The overall background yield is on the other hand smaller than the previous estimates, due to the larger rejection provided by the calorimeter isolation requirement and the loss in acceptance from a few dead optical transmitters of the front-end electronic boards. The observed background yield and the exponential fit to its $m_{\gamma\gamma}$ distribution are used to evaluate the expected sensitivity of this analysis with the current data and to extrapolate it to a luminosity of 1 fb^{-1}. Nuisance parameters (and corresponding Gaussian constraints) are included in the likelihood to account for theoretical uncertainties ($^{+20}_{-15}$%), experimental systematic uncertainties affecting the expected signal yield (luminosity: ±3.4%; photon identification efficiency: ±11%; photon isolation efficiency: ±10%) or invariant mass resolution (±13%). For the signal, the available NNLO cross section calculations are used, since the background is now directly estimated from the data instead of NLO theory calculations. The expected 95% C.L. upper limit on the signal strength for $m_H = 120$ GeV is around 24 with 38 pb^{-1} and around 3.5 with 1 fb^{-1}, as shown in Fig. 4.7. Since no significant peak consistent with the signal resolution is found in the data, an upper limit on μ is set (between 20 and 40) at 95% C.L. as a function of m_H using the first ATLAS data; for $m_H = 120$ GeV, $\mu < 26$.

4.3.3 Higgs boson discovery

After the first study of the 2010 data described in the previous section, ATLAS has frequently updated its $H \to \gamma\gamma$ search, with gradual refinements to the analysis strategy, in order to improve the expected sensitivity to the signal, estimated using improved simulations of the detector response and updated signal cross section estimates [22, 23, 24, 25]. The observed and expected upper limits on μ obtained with the first 1.1 fb^{-1} of data [24], with an ameliorated analysis based on the classification of events in five event categories to improve the sensitivity by about 15% and a more detailed evaluation of the systematic uncertainties, are similar to the expected ones computed in Ref. [20]. In this section I will summarize the results of the analysis of about 11 fb^{-1} of 7 and 8 TeV data [26, 27] that contributed to the joint announcement of the discovery, on July 4th 2012, of a Higgs(-like) boson by the ATLAS and CMS Collaborations.

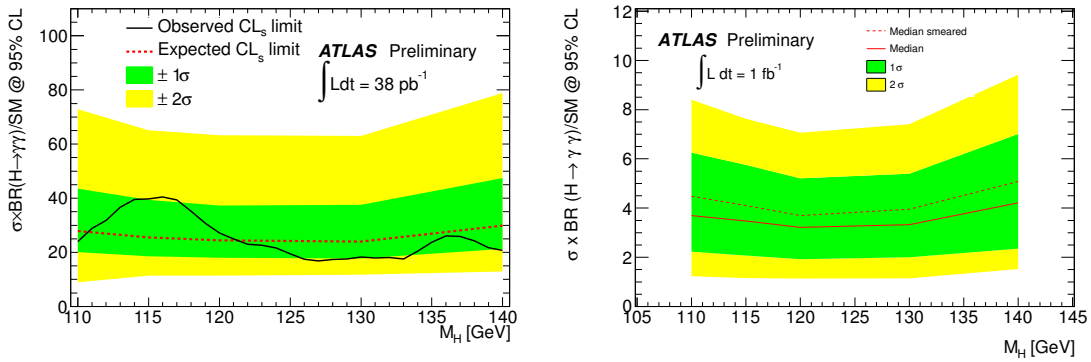

Figure 4.7: Left: expected (dashed line) and observed (solid line) 95% C.L. limit on the signal strength factor set with the CL_s method using 38 pb^{-1} of data collected at 7 TeV [21]. The green (yellow) band correspond to the expected exclusion in the case of a $\pm 1\sigma$ ($\pm 2\sigma$) of the background. Right: estimated upper limit on μ for the same analysis using 1 fb^{-1} of 7 TeV data, with the nominal photon energy resolution (solid line) or using a pessimistic estimate of the resolution (dashed line) [20].

Events are collected using a loose di-photon trigger. The E_{T} threshold applied to each photon candidate is 20 GeV for the 7 TeV data; it is increased to 35 (25) GeV for the leading (sub-leading) photon candidate in the 8 TeV data, to cope with the increased background rates. The efficiency of the trigger is greater than 99% for events passing the final event selection. Only events with good data quality are retained; the corresponding integrated luminosity is 4.8 fb^{-1} at 7 TeV and 5.9 fb^{-1} at 8 TeV, with relative uncertainties of 1.8% and 3.6%, respectively. Events are required to contain at least one reconstructed vertex with two or more associated tracks with $p_{\mathrm{T}} > 0.4$ GeV and two photon candidates in the fiducial pseudorapidity region ($|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$) having $E_{\mathrm{T}} > 40$ GeV and 30 GeV, respectively. Photon candidates are required to have an isolation transverse energy lower than 4 GeV in a cone of radius 0.4 around the photon, and to pass identification criteria based on shower shapes in the electromagnetic calorimeter and on energy leakage into the hadronic calorimeter. For the 7 TeV data, this information is combined in a neural network, tuned to achieve a similar jet rejection as the tight cut-based selection of the previous studies but with higher photon efficiency. For the 8 TeV data, tight cut-based criteria are used to ensure reliable photon performance for recently-recorded data. To compute the di-photon invariant mass, the photon directions are calculated from the position of the event primary vertex and the impact points of the photons in the calorimeter. The event primary vertex is chosen by selecting the vertex which maximises a likelihood based on: the directions of the photons determined using the longitudinal segmentation of the electromagnetic calorimeter (calorimeter pointing), the beam spot, and the $\sum p_{\mathrm{T}}^2$ of the tracks associated to the vertex. For the analysis of the 7 TeV data the reconstructed conversion vertex is also used in the likelihood for photon conversions detected in the silicon detectors of the ID.

About 24 thousand and 35 thousand di-photon candidates are selected in the range $100 < m_{\gamma\gamma} < 160$ GeV in the 7 TeV and 8 TeV data samples, respectively. The relative contributions from $\gamma\gamma$, γj, jj and Drell-Yan events are estimated *in situ* to be approximately 74%, 22%, 3% and 1% using various techniques including a two-dimensional fit to the two photon isolation energies; the 1D projections of the fit and their comparison to the data distributions are shown in Fig. 4.8 for the 8 TeV sample.

To increase the sensitivity to the signal, the events are separated into ten mutually exclusive categories with different invariant mass resolutions and signal-to-background ratios, and a simultaneous maximum likelihood fit to the $m_{\gamma\gamma}$ distributions of the various sub-samples is performed. One category is designed to enrich the fraction ($\approx 70\%$) of signal events produced by VBF, that are characterized by a topology in which the Higgs boson is produced in association with the two scattering quarks that have a large separation in η with small hadronic activity in between. This category requires the presence of at least two anti-k_t jets (with radius parameter $R = 0.4$) with

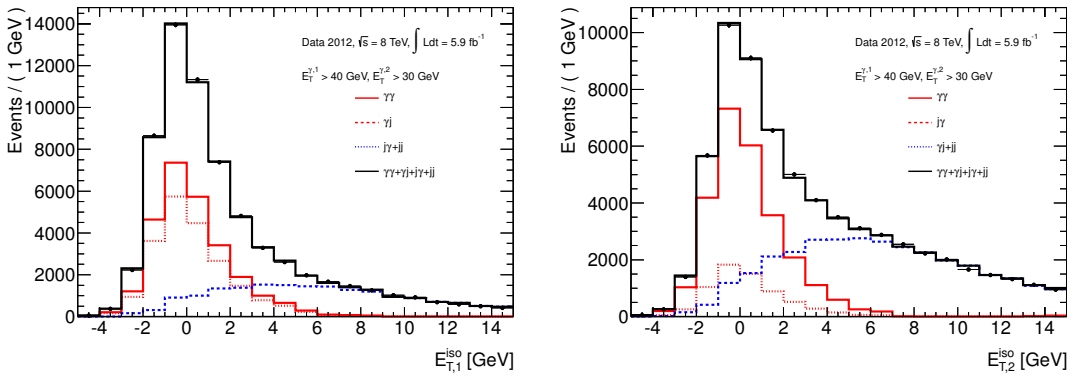

Figure 4.8: Isolation energy distribution of the photon candidates in the selected di-photon sample at 8 TeV (dots with error bars), and projections of the two-dimensional fit (black solid line), for the leading (left) and sub-leading (right) photon candidates. The other lines show the separate contribution from different background sources.

$p_{\mathrm{T}} > 25$ GeV, separated from the photon ($\Delta R_{\gamma j} > 0.4$); to reduce contamination from pile-up jets, the p_{T} threshold is increased to 30 GeV for $|\eta| > 2.5$ at $\sqrt{s} = 8$ TeV, while jets in the ID acceptance ($|\eta| < 2.5$) are required to have JVF>0.75. The leading and sub-leading jet are required to have a large rapidity gap ($|\Delta\eta| > 2.8$), large invariant mass (> 400 GeV), and an azimuthal separation larger than 2.6 between the di-jet and the di-photon system. The other nine categories are defined in the following way: events with both photons reconstructed as unconverted are separated into *unconverted central* ($|\eta| < 0.75$ for both candidates) and *unconverted rest*; the other events are separated into converted central ($|\eta| < 0.75$ for both candidates), converted transition (at least one photon with $1.3 < |\eta| < 1.75$) and converted rest. Each category except the converted transition one is further split in two based on the value of p_{Tt}, the component of the di-photon p_{T} that is orthogonal to the axis defined by the difference between the two photon momenta. This quantity is largely correlated with the di-photon p_{T} but has better detector resolution; signal events, particularly those produced via VBF or associated production, have on average larger p_{Tt} than background events. Events are classified in a *low* p_{Tt} (< 60 GeV) and a *high* p_{Tt} (> 60 GeV) category. The number of data events in each category are given in Table 4.1, together with the expected full-width-at-half-maximum (FWHM) of the signal $m_{\gamma\gamma}$ distribution and S/B at $\sqrt{s} = 8$ TeV. The average FWHM and S/B are 3.9 GeV and 3%, respectively; they vary within $3.2 - 6.1$ GeV and within $1\% - 21\%$ across the categories.

A maximum-likelihood fit to the $m_{\gamma\gamma}$ distribution of the data is performed. In each category, the di-photon invariant mass distribution is fitted with the sum of a signal contribution, described by the sum of a Crystal Ball and a Gaussian functions, and a background contribution, modeled with an analytical function as described below; in the fit, the signal yield is normalized to the SM expectation and the signal strength μ is thus determined.

To study the properties of the expected signal, $H \to \gamma\gamma$ events are generated and fully simulated, for Higgs boson masses between 110 and 150 GeV, in intervals of 5 GeV, at both $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. Gluon-fusion and VBF events are generated with POWHEG [211, 212] (interfaced to PYTHIA for showering and hadronization), while associated production events are generated with PYTHIA. The gluon fusion sample is corrected off-line to take into account the interference with SM $\gamma\gamma$ events (leading to a decrease of the yield by $2 - 5\%$). The parameters of the signal $m_{\gamma\gamma}$ distribution and the signal efficiency ($\approx 40\%$ for $m_H = 125$ GeV) are determined as a function of the Higgs boson mass, in steps of 0.5 GeV, through an interpolation of the values estimated from the simulations, and fixed in the final fit, which is repeated for different m_H hypotheses.

The parameters of the background model, as well as the background yield, are determined directly from the fit to the data. Different background models are chosen for the different categories to achieve the best compromise between limiting the size of a potential bias while retaining good statistical power. The possible bias on the signal yield ("spurious signal") from the choice of the

Table 4.1: Number of events in the data (N_D) and expected number of signal events (N_S) for $m_H = 126.5$ GeV, for each category in the mass range $100-160$ GeV. The mass resolution FWHM and the expected S/B ratio in a mass window around $m_H = 126.5$ GeV that would contain 90% of the expected signal events are also given for the 8 TeV data. The statistical uncertainties on N_S and FWHM are less than 1%.

$\sqrt{s}$	7 TeV		8 TeV			
Category	N_D	N_S	N_D	N_S	FWHM [GeV]	S/B
Unconverted central, low p_{Tt}	2054	10.5	2945	14.2	3.4	0.06
Unconverted central, high p_{Tt}	97	1.5	173	2.5	3.2	0.17
Unconverted rest, low p_{Tt}	7129	21.6	12136	30.9	3.7	0.02
Unconverted rest, high p_{Tt}	444	2.8	785	5.2	3.6	0.07
Converted central, low p_{Tt}	1493	6.7	2015	8.9	3.9	0.04
Converted central, high p_{Tt}	77	1.0	113	1.6	3.5	0.15
Converted rest, low p_{Tt}	8313	21.1	11099	26.9	4.5	0.02
Converted rest, high p_{Tt}	501	2.7	706	4.5	3.9	0.06
Converted transition	3591	9.5	5140	12.8	6.1	0.01
2-jets	89	2.2	139	3.0	3.7	0.21
Total	23788	79.6	35251	110.5	3.9	0.03

background parametrization is estimated through $S+B$ fits to three different sets of high-statistics background-only MC samples, using three different generators (Resbos, Diphox and Sherpa) to model the di-photon background, while the γ-jet and di-jet backgrounds are generated with Sherpa and Pythia, respectively. Detector effects are included in the samples produced with parton-level generators with weighting and smearing techniques. Each of the three MC samples is obtained by mixing the different components in the proportions estimated from data and is normalized to the total number of observed events. Various background pdfs are tested; among those that give a bias smaller than 10% of the expected signal and 20% of the statistical uncertainty on the signal from the background fluctuations for every m_H in the range 110–150 GeV, the one with the smaller number of degrees of freedom is retained. A fourth-order Bernstein polynomial function is thus used for the unconverted rest (low p_{Tt}), converted rest (low p_{Tt}) and inclusive categories, an exponential function of a second-order polynomial for the unconverted central (low p_{Tt})), converted central (low p_{Tt}) and converted transition categories, and an exponential function for all others.

For each category c, the largest absolute signal yield over the m_H range studied, $\sigma_{\mathrm{spurious},c}$, is used as an estimate of the potential bias on the signal yield in data; in the 8 TeV sample it varies between 0.3 and 6.8 events, depending on the category. A possible way to reduce $\sigma_{\mathrm{spurious},c}$ would be to perform the fit in a smaller mass window, at the price of a larger statistical uncertainty; preliminary studies we did of the trade-off between the two components indicated a possible reduction of the total uncertainty on the signal by performing a fit in a narrower $m_{\gamma\gamma}$ window ($\pm 15-20$ GeV) compared to the one used in this measurement (± 30 GeV). For the future measurements at high luminosity, where the systematic uncertainties will be the limiting factor to the accuracy of the results, these studies should be revived and extended and the optimal fit range possibly reconsidered.

Systematic uncertainties due to the uncertainties on the signal efficiency, theoretical cross sections and branching ratios, migrations between categories and invariant mass resolution are accounted for in the fit by introducing nuisance parameters and multiplying the likelihood by the corresponding constraint terms. The uncertainties are treated as fully correlated (same nuisance parameter) between the $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV analyses, with the exception of the systematic uncertainty from the luminosity measurement. The impact of the photon energy scale uncertainty (which translates into a 0.6% uncertainty on the mass scale) is excluded from this procedure, as it was found to spoil the validity of the asymptotic formulae, and evaluated instead using a procedure based on pseudo-experiments. For systematic uncertainties affecting shape parameters or fractions of signal events in the categories, a Gaussian constraint is used: the quantity in the likelihood is

multiplied by a factor $(1+\sigma_p\theta_p)$ and the global likelihood is multiplied by a factor $G(\theta_p)$, where G is a normal distribution centered at zero with unity width, θ_p is the nuisance parameter and σ_p is the relative uncertainty; for uncertainties affecting the expected global yields (luminosity, trigger efficiency, photon identification efficiency, ...), log-normal constraints are usually preferred, in order to avoid the negative tails of the Gaussian distribution: the likelihood is multiplied by $G(\theta)$ and the global observable by $e^{\sigma\theta}$.

Theoretical uncertainties affecting the calculations of the Higgs production cross sections are accounted for by using log-normal constraints for the scale uncertainties and Gaussian constraints for the uncertainties arising from the variations of the PDF set's eigenvalues. The expected yields are then rewritten as:

$$\begin{aligned}
N^{SM}_{\mathrm{gg},c}(\boldsymbol{\theta}) &= N^{\mathrm{SM}}_{\mathrm{gg},c} e^{\sigma_{\mathrm{gg_scale}}\theta_{\mathrm{gg_scale}}}(1+\sigma_{\mathrm{gg_PDF,gg}}\theta_{\mathrm{gg_PDF}}) \\
N^{SM}_{\mathrm{VBF},c}(\boldsymbol{\theta}) &= N^{\mathrm{SM}}_{\mathrm{VBF},c} e^{\sigma_{\mathrm{VBF_scale}}\theta_{\mathrm{VBF_scale}}}(1+\sigma_{\mathrm{qq_PDF,VBF}}\theta_{\mathrm{qq_PDF}}) \\
N^{SM}_{\mathrm{WH},c}(\boldsymbol{\theta}) &= N^{\mathrm{SM}}_{\mathrm{WH},c} e^{\sigma_{\mathrm{WH_scale}}\theta_{\mathrm{WH_scale}}}(1+\sigma_{\mathrm{qq_PDF,WH}}\theta_{\mathrm{qq_PDF}}) \\
N^{SM}_{\mathrm{ZH},c}(\boldsymbol{\theta}) &= N^{\mathrm{SM}}_{\mathrm{ZH},c} e^{\sigma_{\mathrm{ZH_scale}}\theta_{\mathrm{ZH_scale}}}(1+\sigma_{\mathrm{qq_PDF,ZH}}\theta_{\mathrm{qq_PDF}}) \\
N^{SM}_{\mathrm{ttH},c}(\boldsymbol{\theta}) &= N^{\mathrm{SM}}_{\mathrm{ttH},c} e^{\sigma_{\mathrm{ttH_scale}}\theta_{\mathrm{ttH_scale}}}(1+\sigma_{\mathrm{gg_PDF,ttH}}\theta_{\mathrm{gg_PDF}})
\end{aligned} \tag{4.16}$$

and the likelihood is multiplied by

$$G(\theta_{\mathrm{gg\_scale}})G(\theta_{\mathrm{VBF\_scale}})G(\theta_{\mathrm{WH\_scale}})G(\theta_{\mathrm{ZH\_scale}})G(\theta_{\mathrm{ttH\_scale}})G(\theta_{\mathrm{gg\_PDF}})G(\theta_{\mathrm{qq\_PDF}}) \tag{4.17}$$

The nuisance parameters are only seven, to account for five scale uncertainties and five PDF ones, because the PDF uncertainties for the gluon fusion and the $t\bar{t}H$ cross sections are correlated, since both processes are originated by two scattering gluons and their uncertainties are thus originating from the same source, *i.e.* the uncertainty on the gluon PDF $g(x)$; similarly, the PDF uncertainties on the VBF, WH and ZH cross sections are correlated, since all these processes originate from $q\bar{q}$ interactions. Two additional nuisance parameters account for the theoretical uncertainty on the $H\to\gamma\gamma$ branching ratio and the uncertainty on the fraction of gluon fusion+2 jet events in the 2-jet category arising from the choice of the momentum scales in the fixed-order calculations.

The spurious signal systematic uncertainty is included in the model by adding to the expected signal yield μN^{SM}_c, for each of the 10 categories, a term $\sigma_{\mathrm{spurious},c}\theta_{\mathrm{spurious},c}$ and multiplying the likelihood by $G(\theta_{\mathrm{spurious},c})$.

The other experimental systematic uncertainties that are considered are the following. For the overall signal yield, we include contributions from:

- luminosity uncertainty (1.8% at 7 TeV, 3.6% at 8 TeV);
- trigger, photon reconstruction and identification efficiency (9−12%, mostly due to the photon identification);
- efficiency of the photon isolation requirement (0.4%);
- impact of the photon energy scale uncertainty on the selection efficiency (0.3%);

For migrations of signal events between different categories, we evaluate the effects of

- the Higgs boson p_{T} modeling in the simulation (1.1% for low p_{Tt} categories, 12.5% for high p_{Tt} ones, 9% in the 2-jet category). It is estimated by varying the scales and PDFs within their uncertainties in the program HQT2 that is used to tune the Higgs boson p_{T} distribution in POWHEG.
- the underlying-event model (relevant only for the 2-jet category, it is 6% on the contribution from VBF events and 30% on the contribution of other processes). It is estimated using alternative underlying-event tunes in the simulation.
- detector material uncertainty ($\approx 4\%$ for the nine categories based on the photon conversion status), estimated using simulations with a different detector model.
- pile-up effects ($2-3\%$), estimated by comparing signal fractions in the various categories in low or high pile-up simulated events.

- jet energy scale (19% for the 2-jet category and $< 4\%$ for the others), estimated by varying the scale within uncertainties in the simulation.
- JVF requirement (13% on the 2-jet category, 8 TeV data only), from a data/MC efficiency comparison in Z+2 jets events.

For the signal mass resolution, the relative uncertainties arising from the uncertainty on the calorimeter energy resolution (12% on $\sigma_{m_{\gamma\gamma}}$), on the electron-to-photon extrapolation of the energy scale (6%) and on the pile-up modeling in the simulation (4%) add up to a relative uncertainty of 14%.

The total number of nuisance parameters $\boldsymbol{\theta}$, *i.e.* the parameters floating in the fit together with μ, is thus 88: 9 for theory uncertainties, 11 for experimental uncertainties, 10 for the spurious signal terms, 20 for the background yields, and 38 for the background shape parameters.

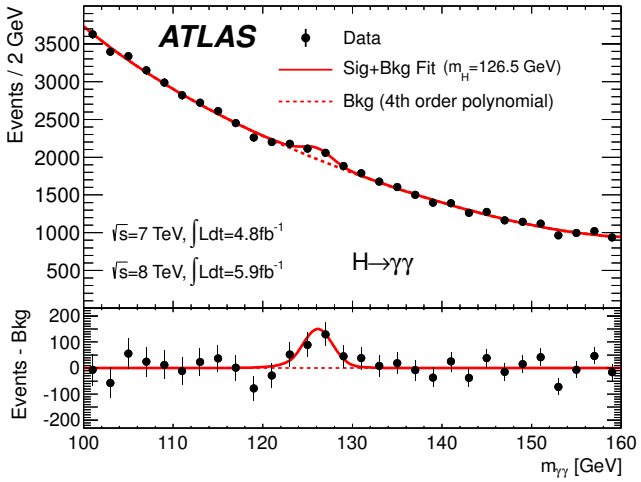

Figure 4.9: Invariant mass distribution of di-photon candidates for the combined $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data samples. The results of a fit to the data of the sum of a signal component fixed to $m_H = 126.5$ GeV and a background component described by a fourth-order Bernstein polynomial, and of a background-only fit, are superimposed. The bottom inset displays the residuals of the data with respect to the fitted background component.

The distribution of the di-photon invariant mass of the selected events, summed over all categories, is shown in Fig. 4.9. The results of a fit to the data of the sum of a signal component fixed to $m_H = 126.5$ GeV and a background component described by a fourth-order Bernstein polynomial, and of a background-only fit, are superimposed. An excess of events over the background-only expectation, consistent with the expected signal resolution at $m_H = 126.5$ GeV, is clearly visible.

Using the statistical techniques described in Sec. 4.2, the following results have been obtained:

- the p_0 (Fig. 4.10, top-left), which quantifies the compatibility of the data with the background-only hypothesis, evaluated as a function of m_H, is minimum (2×10^{-6}) at $m_H = 126.5$ GeV and corresponds to a (local) significance of 4.5σ. After correcting for the trial factor [213], *i.e.* the ratio between the probability of observing the excess at some fixed mass point, to the probability of observing it anywhere in the tested range, the significance of the excess is 3.6σ. The expected p_0 at the same mass is 7×10^{-3} (2.4σ significance).
- a SM Higgs boson is excluded at 95% C.L. for $112 < m_H < 122.5$ GeV and for $132 < m_H < 143$ GeV, while the expected exclusion range in the background-only hypothesis is $110 < m_H < 139.5$ GeV (Fig. 4.10, top-right).
- the best-fit signal strength (Fig. 4.10, bottom-left) for $m_H = 126.5$ GeV is $\hat{\mu} = 1.8 \pm 0.5$, with similar contributions to the total uncertainty from the statistical, the experimental systematic and the theoretical systematic components.

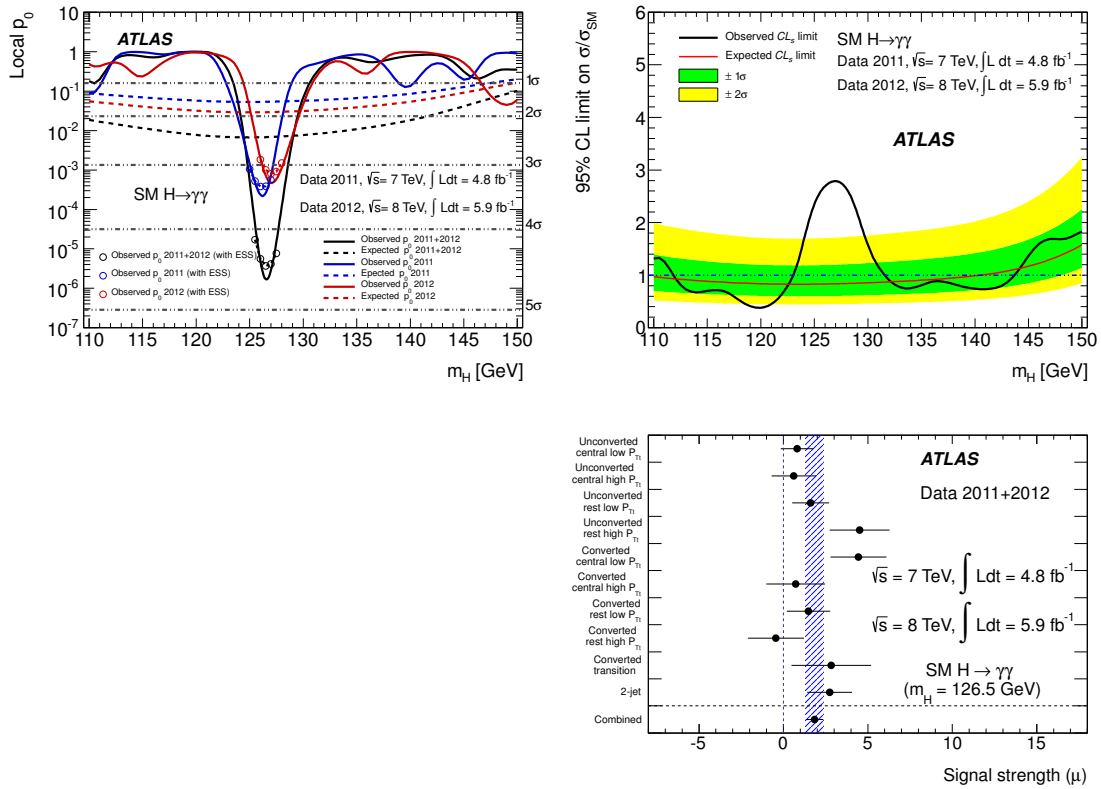

Figure 4.10: Results of the $H \to \gamma\gamma$ search based on ≈ 11 fb^{-1} of 7 and 8 TeV data [27]. Top-left: expected and observed local p_0 values for a SM Higgs boson as a function of the hypothetical Higgs boson mass (m_H). The observed p_0 including the effect of the photon energy-scale uncertainty on the mass position is included via pseudo-experiments and shown as open circles. Top-right: expected and observed CL_s limit on the signal strength as a function of the assumed Higgs boson mass. The dark (green) and light (yellow) bands indicate the expected limits with $\pm 1\sigma$ and $\pm 2\sigma$ fluctuations, respectively. Bottom-left: best-fit value for the signal strength as a function of the assumed Higgs boson mass. Bottom-right: best-fit value for the signal strength in the different categories at $m_H = 126.5$ GeV. The blue band corresponds to the error of the combined result.

- when allowing each category to have its own signal strength, the fit returns consistent values of μ across all the categories (Fig. 4.10, bottom-right).

As a cross-check, the analysis is repeated without dividing the dataset into categories; the observed local significance is reduced to 3.5σ. From the results presented here and the simultaneous observation of an excess of events in the $H \to ZZ^{(*)} \to 4\ell$ channel for $m_{4\ell} \approx 125$ GeV, with a local significance of 3.6σ, and of a broad excess of events in the same mass region in the $H \to WW^{(*)} \to \ell\nu\ell\nu$ channel, with a local significance of 2.8σ, both consistent with the expectations for a SM Higgs boson of mass ≈ 126 GeV, the observation of a new particle with mass near 126 GeV and in agreement with the SM Higgs boson hypothesis ($\hat{\mu} = 1.4 \pm 0.3$) can be established, with a global combined significance of 5.1σ [27].

4.4 Search for the SM Higgs boson decaying to a photon and a Z boson, $H \to Z\gamma$

Once the observation of a new particle consistent with the SM Higgs hypothesis is established, it becomes fundamental, in order to discriminate between the SM and alternative theories, to fully characterize the properties of this particle, including parity, spin, and decays to as many final states as possible. After the summer 2012 I have thus started to search for the decay $H \to Z\gamma$, $Z \to \ell^+\ell^-$ ($\ell = e$ or μ), using the full 7 and 8 TeV pp ATLAS data; the first public results of that work, available at the beginning of March 2013 and documented in Ref. [28], are summarized here.

Measurements of (or limits on) the $H \to Z\gamma$ decay rate can provide insight into models beyond the SM. The decay rate can help determine whether the new boson is the Higgs boson or a member of other electroweak singlets or triplets. Moreover, because the $H \to Z\gamma$ decay proceeds via electroweak loop coupling to the Higgs boson, it can provide direct evidence of new heavy particles running in the loop, and comparing the $H \to Z\gamma$ and $H \to \gamma\gamma$ rates can yield hints about the quantum numbers of these new particles. Some models [46, 47, 48, 49] predict significant enhancements over the SM rate; however, many of these alternative hypotheses are already severely limited by the known $H \to \gamma\gamma$, $H \to ZZ^*$, and SM $Z\gamma$ measurements [214].

For a SM Higgs boson mass of 125 GeV, the predicted $pp \to H \to Z\gamma \to \ell\ell\gamma$ cross section, including $BR(Z \to \ell^+\ell^-) = 6.7\%$ [215], is 2.3 (1.8) fb at $\sqrt{s} = 8$ (7) TeV, roughly similar to that of $pp \to H \to ZZ^* \to 4\ell$ and only 5% of that of $pp \to H \to \gamma\gamma$; as a consequence, if the efficiency is similar to that for $H \to 4\ell$ or $H \to \gamma\gamma$ events ($\varepsilon \approx 30\%$), we expect to reconstruct less than 20 signal events in the SM hypothesis, and possibly many more in case of a non-SM enhancement of their rate. The $\ell\ell\gamma$ final state can be produced also from internal photon conversion in Higgs-boson decays to di-photons ($H \to \gamma^*\gamma \to \ell\ell\gamma$) or from radiation emission by leptons in Higgs-boson decays to di-leptons ($H \to \ell\ell^* \to \ell\ell\gamma$), collectively called in the following Dalitz decays [216]; our selection purposedly suppresses these additional $H \to \ell\ell\gamma$ decay processes, as they are not included in the SM theoretical calculations to which we compare our observed yields.

Similarly to the $H \to \gamma\gamma$ decays, the main backgrounds originate from (i) irreducible SM Z+γ events with a true prompt photon, and (ii) reducible Z+jet events due to misidentification of a jet as a photon. Much smaller contributions are expected from other processes ($t\bar{t}$ and W/Z). The irreducible background originates either from diboson production in the t, u channels, from final-state-radiation (FSR) in radiative Z boson decays ($Z \to \ell\ell\gamma$), or from parton-to-photon fragmentation. Although the background level for $H \to Z\gamma$ is reduced compared to $H \to \gamma\gamma$, it is orders of magnitude higher than that for $H \to ZZ^* \to 4\ell$ and thus the sensitivity of this channel to a SM Higgs boson is anticipated to be small.

Unless explicitly specified, the selection criteria at $\sqrt{s} = 8$ TeV and $\sqrt{s} = 7$ TeV are identical. Values quoted in parentheses correspond to the $\sqrt{s} = 7$ TeV run period. The efficiency of the selection and the properties of the selected signal events are studied using simulated samples of $H \to Z\gamma$ decays generated with Powheg interfaced to Pythia 8.170, using the CT10 PDFs. Signal events are generated for Higgs boson masses between 120 and 150 GeV, in intervals of 5 GeV, at both $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. The simulation is corrected to take into account known data-MC differences on photon and lepton efficiencies and energy or momentum resolution.

Events are collected using the lowest threshold, unprescaled single-lepton or di-lepton triggers

in data with good quality. For the single-muon trigger the transverse momentum threshold is 24 (18) GeV, while for the single-electron trigger the transverse energy threshold is 24 (20) GeV. For the di-muon triggers the thresholds are $p_{\mathrm{T}} > 13$ (10) GeV for each muon, while for the di-electron triggers the thresholds are $E_{\mathrm{T}} > 12$ GeV for each electron. At $\sqrt{s} = 8$ TeV an asymmetric di-muon trigger is also used with $p_{\mathrm{T1}} > 18$ GeV and $p_{\mathrm{T2}} > 8$ GeV. The trigger efficiency with respect to signal events passing the selection criteria is around 99% in the $ee\gamma$ channel and 92% in the $\mu\mu\gamma$ channel due to the reduced geometric acceptance of the muon trigger system in the $|\eta| < 1.05$ region. The resulting integrated luminosity corresponds to 20.7 fb^{-1} (4.6 fb^{-1}), with a relative uncertainty of 3.6% (1.8%) at $\sqrt{s} = 8$ TeV ($\sqrt{s} = 7$ TeV) [217, 70].

Muon candidates are reconstructed as described in Sec. 1.5. All muon candidates identified using the information from the MS are required to have transverse momentum $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.7$, while the ones that are tagged by the calorimeters must have $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 0.1$. The inner detector tracks associated to muons that are identified inside the ID acceptance are required to have a minimum number of associated hits in each of the ID sub-detectors (to ensure good track reconstruction) and to have transverse (longitudinal) impact parameter d_0 (z_0) smaller than 1 mm (10 mm) with respect to the hard-scattering primary vertex, defined as the primary vertex with the largest sum of the squared transverse momenta of the tracks associated to it.

Electron candidates are required to have a transverse energy greater than 10 GeV and pseudorapidity $|\eta| < 2.47$, and to pass loose requirements on the quality of the reconstructed track and on the longitudinal and transverse shower profiles of the energy cluster in the electromagnetic calorimeter. To suppress non-prompt electron candidates, the electron tracks are required to have a longitudinal impact parameter, with respect to the primary vertex, smaller than 10 mm and to have a hit in the b-layer when passing through an active b-layer module.

Photon candidates are required to have transverse energy greater than 15 GeV, $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$, to pass tight identification requirements on the shower shapes measured in the first two longitudinal layers of the electromagnetic calorimeter and on the leakage in the hadronic calorimeter, and to have an isolation transverse energy in the calorimeter of less than 4 GeV. The isolation requirements suppresses the Z+jet background as well as the fragmentation component in $Z+\gamma$ events. In case of multiple photon candidates, the one with largest transverse energy is used to reconstruct the decay of the Higgs boson candidate.

An overlap removal between electrons and muons that pass all selection criteria and share the same inner detector track within a cone of $\Delta R < 0.02$ is performed: if the muon is identified by the MS, then the electron candidate is discarded, otherwise the muon candidate is rejected. Photon candidates that are within $\Delta R < 0.3$ of a selected electron or muon candidate are also rejected, thus suppressing background from FSR events.

Z boson candidates are reconstructed from pairs of same-flavor, opposite-sign lepton (e or μ) candidates. Muon pairs are required to have no more than one muon identified either without the MS or in the region outside the ID acceptance. Electron candidates are required to pass medium identification criteria. In case of multiple Z candidates in the same event, the candidate with the invariant mass closest to the Z pole is selected. To suppress events from FSR $Z \to \ell\ell\gamma$, the invariant mass of the selected di-lepton pair must be larger than the PDG value of the Z boson mass minus 10 GeV; this requirement also reduces the contribution to the signal from internal photon conversions in $H \to \gamma\gamma$ to a negligible level [218]. The two leptons from the Z boson decay must also pass track and calorimeter isolation requirements. The normalized track isolation, *i.e.* the sum of the transverse momenta of tracks inside a cone of $\Delta R < 0.2$ around the lepton (excluding the lepton track) divided by the lepton p_{T}, must be smaller than 0.15. For electrons, the normalized calorimetric isolation $E_{\mathrm{T}}^{\mathrm{iso}}/E_{\mathrm{T}}$ in a cone of radius 0.2 around the electron must be lower than 0.2, where $E_{\mathrm{T}}^{\mathrm{iso}}$ is computed in the same way as for photons. For muons, the normalized calorimetric isolation $E_{\mathrm{T}}^{\mathrm{iso}}/p_{\mathrm{T}}$ is required to be less than 0.3 (0.15 in case of muons without an ID track), where $E_{\mathrm{T}}^{\mathrm{iso}}$ is computed from the sum of the transverse energy of the calorimeter cells inside a cone of $\Delta R < 0.2$ around the muon direction, excluding the small energy deposited by the muon itself. For both the track- and calorimeter-based isolation any contributions arising from the other lepton from the candidate Z decay are subtracted. Additional selections are also applied on the track impact parameter significance of the leptons from the Z candidate decay: the transverse impact parameter significance $|d_0|/\sigma_{d_0}$ of the ID track associated to leptons within the acceptance of the inner detector is required to be less than 3.5 for muons and 6.5 for electrons,

whose distribution is broader because of bremsstrahlung.

Higgs boson candidates are reconstructed from the combination of the photon candidate with the two leptons from the Z boson candidate. To improve the three-body $\ell\ell\gamma$ invariant mass resolution and thus improve discrimination against non-resonant background events, two corrections are applied to the three-body mass $m_{\ell\ell\gamma}$: (i) the photon pseudorapidity η^γ and its transverse energy $E_{\rm T}^\gamma = E^\gamma / \cosh \eta^\gamma$ are recalculated from the positions of the primary vertex and of the photon impact point in the calorimeter, and (ii) the lepton four-momenta are recomputed by means of a Z-mass constrained kinematic fit previously used in the $H \to 4\ell$ search [27]. After corrections, the expected core resolution of the $m_{\ell\ell\gamma}$ distribution is 1.6 GeV and the FWHM is 4 GeV for $m_H = 125$ GeV.

After applying all the selection criteria, the number of $Z\gamma$ candidates in the $\sqrt{s} = 8$ (7) TeV data sample is 13978 (1927) in the $Z \to ee$ channel and 16678 (2621) in the $Z \to \mu\mu$ channel. The expected signal efficiency for $m_H = 125$ GeV is 24.6% (20.4%) for $Z \to ee$ and 29.7% (26.5%) for $Z \to \mu\mu$, where the increase at 8 TeV is largely due to the increased photon efficiency from the reoptimised photon identification criteria; the corresponding expected signal yields, neglecting the contribution from the Dalitz $H \to \ell\ell\gamma$ decays, are 5.9 (0.9) and 7.2 (1.1). For $120 < m_H < 150$ GeV, the total expected signal yield varies between 10 (for $m_H = 120$ GeV) and 24 (for $m_H = 140$ GeV).

The composition of the selected sample in terms of the various background processes is measured *in situ* using the same two-dimensional sideband technique developed for the prompt photon cross section measurements (Sec. 3.3.2), based on the distribution of the photon identification and isolation variables in control regions enriched in Z+jets events, after subtracting from the data the contribution from the $t\bar{t}$ and WZ backgrounds, that are estimated from the simulation using the NLO MC cross sections, on which a conservative uncertainty of $\pm 50\%$ is applied [1]. In the study, the Z+γ contamination in the Z+jet background control regions and the correlation between the photon identification and isolation variables for Z+jet events are determined from simulated events: SM Z+γ events with up to three additional partons in the LO matrix element are generated with Sherpa 1.4.1 (1.4.0), using the CT10 PDFs, while Drell-Yan events with up to 5 partons produced with the Z boson are generated with Sherpa 1.4.0 (using the CT10 PDFs) and with Alpgen 2.13 (using the CTEQ6L1 PDFs) interfaced to Herwig 6.510 for parton shower and fragmentation into particles, to Jimmy 4.31 [148] to model the underlying event, and to Photos [221] for final state QED radiation. The resulting fractions of Z+γ, Z+jets and other backgrounds are around 82%, 17% and 1% at both $\sqrt{s} = 7$ and 8 TeV. The uncertainty on the Z+γ purity is around 5%, dominated by the uncertainty on the correlation between the photon identification and isolation in Z+jet events, which is estimated by comparing the Alpgen and Sherpa predictions. A comparison between the three-body invariant mass distribution in $\sqrt{s} = 8$ TeV data and simulation, after scaling the MC background yields to the values determined in data, is shown in Fig. 4.11. A good agreement between data and simulation is observed in the distributions of $m_{\ell\ell\gamma}$, as well as in the distributions of several other kinematic quantities that have also been studied.

Similarly to the $H \to \gamma\gamma$ case, the search for the $H \to Z\gamma$ signal is performed through an unbinned maximum-likelihood signal+background fit to the distribution of a discriminating variable. A simultaneous fit is performed to the four orthogonal categories defined by the lepton flavor (e or μ) and the center-of-mass energy (7 or 8 TeV) of the dataset. Instead of the three-body invariant mass of the final state particles, $m_{\ell\ell\gamma}$, we choose as discriminating variable the difference between the three-body and the di-lepton invariant masses, $\Delta m = m_{\ell\ell\gamma} - m_{\ell\ell}$. The overall background distribution is expected to be smooth for both these variables and thus, like for $H \to \gamma\gamma$, a simple analytical shape whose parameters are directly fitted on data can be used to represent the background. However, a fit to the Δm distribution has the advantage that a possible residual signal from FSR in $H \to \mu\mu$ decays [2] will not bias the fitted μ, which is extracted by normalizing the fitted yield to theoretical predictions which do not include yet the Dalitz Higgs decay amplitudes.

The Δm distribution of signal events is well described empirically by the sum of a Crystal Ball

1. $t\bar{t}$ events are generated with MC@NLO [219, 220] interfaced to Herwig 6.510 for parton shower and fragmentation and to Jimmy 4.31 [148] for the underlying event, using the CT10 PDFs. WZ events at $\sqrt{s} = 7$ TeV are generated with up to three additional partons from the matrix element using Sherpa 1.3.1 and the CTEQ6L1 PDFs. At $\sqrt{s} = 8$ TeV, WZ events are generated with Powheg interfaced to Pythia 8.165.

2. $H \to ee\gamma$ from FSR in $H \to ee$ in the SM is expected to be negligible due to the small electron mass

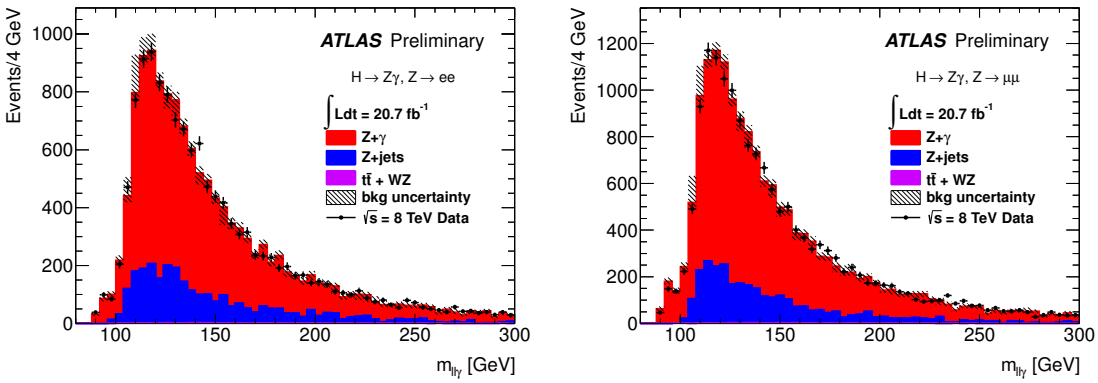

Figure 4.11: Three-body invariant mass ($m_{\ell\ell\gamma}$) distribution of selected events in $\sqrt{s} = 8$ TeV data (dots) and from the various background sources (histograms, from the simulation) normalized to the yields determined as described in the text, for $Z \to ee$ (left) and $Z \to \mu\mu$ (right) channels. The background uncertainty includes statistical uncertainties and systematic uncertainties from the inputs taken from the simulation, as detailed in the text.

lineshape and a small wide Gaussian component, as shown in Fig. 4.12. The core resolution is 1.6 GeV and the FWHM is 4 GeV; the width of the distribution is dominated by the Z-boson intrinsic width, with a sub-leading contribution of ≈ 1 GeV from the photon energy resolution (which is around 3%).

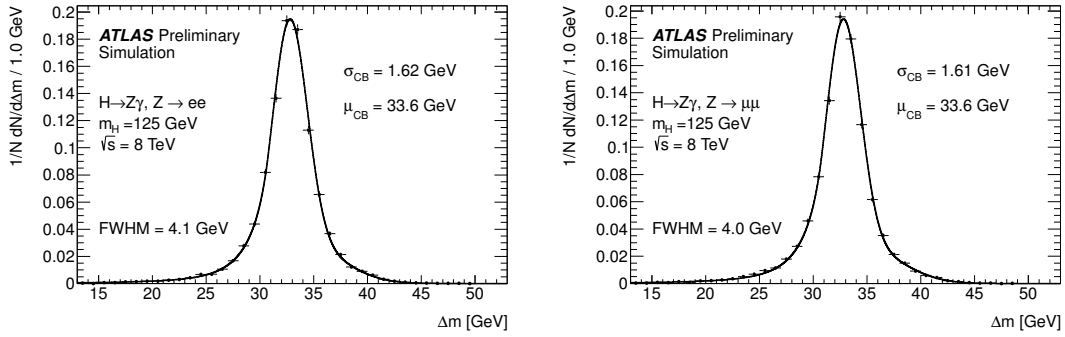

Figure 4.12: Distribution (normalized to unit area) of the difference Δm between the final state three-body invariant mass $m_{\ell\ell\gamma}$ and the di-lepton invariant mass $m_{\ell\ell}$ for signal events passing the full selection (dots), for $m_H = 125$ GeV and $\sqrt{s} = 8$ TeV. The line overlaid represents the fit of the distribution with a model composed of the sum of a Crystal Ball (CB) and a Gaussian (GA) function. Left: electron channel, right: muon channel.

The background model is chosen, among several functional forms including polynomials of various orders, as well as non-polynomial functions such as exponential, Crystal Ball, Crystal Ball+Gaussian, and Crystal Ball+Landau distributions, using the same criteria applied for the $H \to \gamma\gamma$ analysis, *i.e.* by keeping the model with the lowest number of degrees of freedom and yielding a small-enough fitted signal in signal+background fits to high-statistics simulated background-only samples. The chosen background model is a third-order Chebychev polynomial in the fit range $24 < \Delta m < 64$ GeV, and the spurious signal is treated as a systematic uncertainty on the fitted signal yield . Fig. 4.13 shows the results of background-only fits to the data in the two lepton-flavor categories for the $\sqrt{s} = 8$ TeV data.

The systematic uncertainties are accounted for, as in the $H \to \gamma\gamma$ search, by introducing nuisance parameters in the likelihood, which is then multiplied by the corresponding constraint terms. Theoretical uncertainties on the production cross section (from the choice of the energy

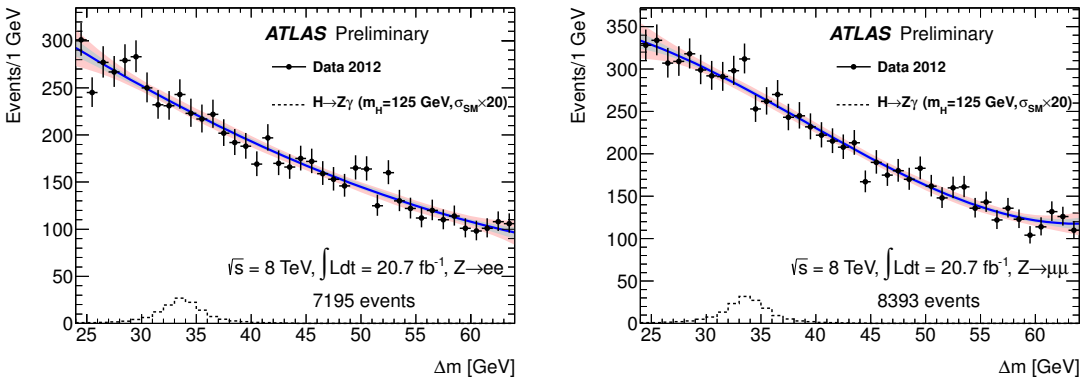

Figure 4.13: Background-only fits to the distribution of the mass difference Δm of selected events in data, for $Z \to ee$ (left) and $Z \to \mu\mu$ (right), at $\sqrt{s} = 8$ TeV. A third order polynomial is used for the fit. Dots correspond to data, the blue line is the fit result and the gray and light red bands are the 1σ and 2σ uncertainty bands from the statistical uncertainties on the fitted background model parameters. The dashed histograms correspond to the SM signal expectation, for a Higgs boson mass of 125 GeV, scaled by a factor 20 for clarity.

scales used for the fixed-order calculation and the uncertainties on the PDF eigenvalues and on the value of α_s) and on the $H \to Z\gamma$ branching ratio are taken from Refs. [42, 43]. The main sources of experimental systematic uncertainties and their contributions to the $H \to Z\gamma$ expected signal yields and parameters of the signal Δm distributions are listed in Table 4.2 for $m_H = 125$ GeV. All systematic uncertainties, except that on the luminosity, are treated as correlated between the $\sqrt{s} = 7$ TeV and the $\sqrt{s} = 8$ TeV data. The experimental systematic uncertainties on the expected

Table 4.2: Summary of the systematic uncertainties on the signal yield and invariant mass distribution for $m_H = 125$ GeV, at $\sqrt{s} = 8(7)$ TeV.

Systematic Uncertainty	$H \to Z(ee)\gamma(\%)$	$H \to Z(\mu\mu)\gamma(\%)$
Signal Yield		
Luminosity	3.6 (1.8)	3.6 (1.8)
Trigger efficiency	0.4 (0.2)	0.8 (0.7)
Acceptance of kinematic selection	4.0 (4.0)	4.0 (4.0)
γ identification efficiency	2.9 (2.9)	2.9 (2.9)
electron reconstruction and identification efficiency	2.7 (3.0)	
μ reconstruction and identification efficiency		0.6 (0.7)
e/γ energy scale	1.4 (0.3)	0.3 (0.2)
e/γ isolation	0.4 (0.3)	0.4 (0.2)
e/γ energy resolution	0.2 (0.2)	0.0 (0.0)
μ momentum scale		0.1 (0.1)
μ momentum resolution		0.0 (0.1)
Signal Δm resolution		
e/γ energy resolution	5.0 (5.0)	2.4 (2.4)
μ momentum resolution		0.0 (1.5)
Signal Δm peak position		
e/γ energy scale	0.2 (0.2) GeV	0.2 (0.2) GeV
μ momentum scale		negligible

signal yields have been evaluated as follows:

- acceptance of the kinematic requirements: the acceptances estimated with simulated signal events generated either using Powheg or Mcfm, both interfaced to Pythia, are compared.
- photon identification efficiency: at $\sqrt{s} = 7$ TeV, the signal yield is recomputed by varying

the photon identification efficiency scale factors within their uncertainties and the relative variation is considered as a systematic uncertainty. At $\sqrt{s} = 8$ TeV a conservative estimate of the uncertainty on the photon identification efficiency obtained from a comparison between data-driven measurements and the simulated efficiencies is used. This amounts to 2.5% for $E_{\mathrm{T}} < 40$ GeV and for unconverted photons with $|\eta| > 1.81$, and to 1.5% otherwise.

- photon and electron calorimeter isolation requirements: the signal efficiency is recomputed by shifting, in the simulation, the photon and electron calorimeter isolation energies by the average difference (≈ 100 MeV) observed between the isolation energy distributions in data and Monte Carlo of photons and electrons, selected either in di-photon enriched events or in a control sample of electrons from $Z \to ee$.
- photon and electron energy scales: the electromagnetic energy scale corrections (applied to the data) are varied within their uncertainties in the simulation and the relative variation in the predicted signal yield is considered as a systematic uncertainty.
- photon and electron energy resolution: the electromagnetic energy smearing correction applied to the simulation is varied within its uncertainty.
- muon momentum scale and resolution: the muon momentum corrections applied to the simulation are varied within their uncertainties.
- lepton trigger, reconstruction and identification efficiency: the efficiency scale factors applied to the simulation are varied within their uncertainties.

Other sources of uncertainties (the efficiency of the impact parameter and track isolation requirements for the leptons and of the calorimeter isolation selection for muons) have been estimated comparing the efficiencies in data and Monte Carlo for control samples of leptons from Z decays and found to be negligible. The total relative uncertainty on the signal efficiency is around 5%. The experimental systematic uncertainties on the signal Δm peak position are evaluated as the shift of the peak when recomputing the signal Δm distribution after varying either the electromagnetic (e/γ) energy scale or the muon momentum scale within their uncertainties. The experimental systematic uncertainties on the signal Δm resolution are evaluated as the relative variation of the width of the signal Δm distribution after varying either the electromagnetic energy or muon momentum smearing corrections within their uncertainties. The final likelihood includes 8 nuisance parameters describing theory uncertainties, 13 accounting for the experimental uncertainties, 2 for the spurious signal term, 4 for the background yields and 12 for the background shape parameters; the total number of nuisance parameters is thus 39.

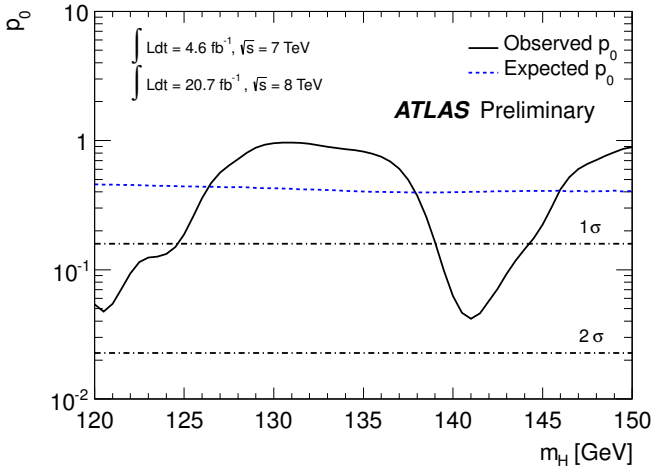

Figure 4.14: Expected (dashed blue line) and observed (solid black line) p_0 (compatibility of the data with the background-only hypothesis) as a function of the Higgs boson mass, using 4.6 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV and 20.7 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV.

The expected and observed p_0 values are shown in Fig. 4.14 as a function of the Higgs boson mass. The expected p_0 ranges between 0.40 and 0.46 for $120 < m_H < 150$ GeV, corresponding to local significances around 0.25 σ. The observed p_0 distribution is compatible with the data being composed of background only. The smallest p_0 (0.042), corresponding to the largest deviation of the data from the background-only hypothesis, with a local significance of 1.61 σ, occurs for a mass of 141 GeV. The expected p_0 at $m_H = 125$ GeV is 0.443, corresponding to a local significance of 0.14 σ, while the observed one is 0.188 (0.89 σ).

Upper limits on the production cross section of $H \to Z\gamma$ are set at 95% C.L. with the CL_s technique. Both observed limits, computed using real data, and expected limits, computed using an Asimov dataset generated in the $\mu = 0$ hypothesis, are shown in Fig. 4.15. The expected 95% C.L. limit ranges between 7.3 and 22 times the Standard Model, and the observed one varies between 5.4 and 37 times the Standard Model, for a Higgs boson mass between 120 and 150 GeV. In particular, for a mass of 125 GeV, consistent with the mass of the recently discovered Higgs-like boson, the expected and observed limits are equal to 13.5 and 18.2 times the Standard Model, respectively. The results are dominated by the statistical uncertainties: neglecting all systematic uncertainties, the observed (expected) 95% C.L. limit on the cross section at 125 GeV is 17.4 (12.9) times the Standard Model prediction.

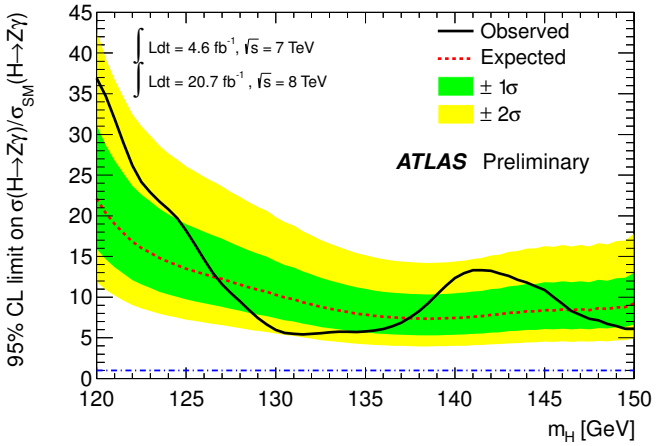

Figure 4.15: Observed 95% C.L. limits (solid black line) on the production cross section of a SM Higgs boson decaying to $Z\gamma$, as a function of the Higgs boson mass, using 4.6 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV and 20.7 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV. The median expected 95% C.L. exclusion limits (dashed red line) are also shown. The green and yellow bands correspond to the $\pm 1\sigma$ and $\pm 2\sigma$ intervals.

The sensitivity of the current analysis is still rather far from the expected SM rate. However, work is ongoing to improve the analysis, by classifying the events in (a few) categories with different S/B and enhancing the selection efficiency. Preliminary studies show that the expected upper limit on μ with the current data could decrease to ≈ 9, which implies that in the next run of the LHC – thanks to the larger ($2x$) cross sections at 14 TeV and the order of magnitude increase in integrated luminosity – $H \to Z\gamma$ decays could be observed for $\mu = 1$ or for a larger μ, similar to that measured by ATLAS in the $H \to \gamma\gamma$ channel, thus providing additional useful information on the particle that was recently discovered.

Conclusion

In this document, prepared for my Habilitation à Diriger des Recherches, I have summarized a part of the research activities that I carried on during the past four years and a half, after I left the BABAR group in Pisa – at the end of my third year of Post.Doc. – and joined the ATLAS group of the Laboratoire de Physique Nucleaire et des Hautes Energies in Paris. The common underlying trait of the data analysis activities described here is the experimental investigation of processes producing prompt photons in the final states of the *pp* collisions collected by ATLAS at the LHC.

I began working on this topic at the beginning of 2009, mainly motivated by the expertise and the involvement of the LPNHE ATLAS group in the construction and commissioning of the liquid-argon electromagnetic calorimeter and in the reconstruction of electrons and photons, and by the undeniable charm (at least for me!) of the Higgs-boson physics, with the hope to contribute some day to the discovery of the Higgs boson by searching for its decay to the di-photon final state. This is why one of the first topics I contributed to was the MC-based estimation of the sensitivity of the ATLAS $H \to \gamma\gamma$ search (Sec. 4.3.1) and the (re-)optimization of the photon identification criteria (Sec. 2.1) before the start of the data-taking; the $H \to \gamma\gamma$ search has then been a constant interest of mine, and I contributed to the various updates of the analysis (Sec. 4.3.2 and Sec. 4.3.3) that finally led to the *discovery* of a particle consistent with the SM Higgs-boson predictions.

Over the years, however, I have learnt that the $H \to \gamma\gamma$ search, as important as it certainly is, is not the only interesting topic worth focusing on while analysing ATLAS data containing reconstructed photon candidates.

One reason is that to positively identify the particle recently discovered as the Higgs boson of the Standard Model we need to characterize as accurately as possible its properties, including its spin, its parity, and its coupling to other particles, through measurements of various final states sensitive to these different properties and comparing them with the SM expectations. We thus need to expand the set of final states that have been investigated so far to other decay channels with non-negligible branching ratios that could be possibly enhanced in BSM models and provide hints on an extended or non-standard Higgs sector: to this purpose I have led the first ATLAS search of the $H \to Z\gamma$ decay, based on the full data collected until the end of 2012 (Sec. 4.4). In parallel, we need to determine as accurately as possible both the experimental and theoretical inputs used in the measurement or calculation of the Higgs boson effective production cross sections: that is why I was deeply involved in the *in situ* measurement of the photon identification (Sec. 2.2) and trigger (Sec. 2.4) efficiencies and in the measurements of prompt photon production cross sections that have the potential to reduce the theoretical uncertainties on the gluon PDF in the proton and thus the large (indeed, dominating) PDF uncertainties on the Higgs boson production cross section (Sec. 3.3–3.6). As these searches target final states that are affected by significantly larger backgrounds, the selection criteria must be carefully optimized and a good understanding of the background composition is important, though sometimes not needed in the final estimation of the signal yield, for tuning the selection to achieve an optimal background rejection and for improving the reliability of the simulation programs used to study the properties of the background itself. The measurement I did of the SM photon-jet and di-photon cross sections (Sec. 3.5 and 3.7) are the results of a personal effort to measure the cross sections of the main backgrounds to the $H \to \gamma\gamma$ search, developing techniques to discriminate between them and the di-jet background, and to compare the results to both LO parton-shower and (N)NLO parton-level event generators.

A second reason for not limiting myself to the $H \to \gamma\gamma$ search is that the physics reach of ATLAS using photons is simply not confined to the Higgs sector, for various motives. An obvious

one, at least until the end of 2011, is that the existence of an Higgs boson candidate was not granted: indeed, when I started most of the activities described here, it was still unproven. High-energy photons at the LHC can thus be a portal for exploring alternative models of physics beyond the Standard Model, showing up for instance in the searches of exotice photon-jet or di-photon resonances, which also benefit from the work I did on the improvement and of the data-driven measurement of the photon selection performances in the ATLAS detector. More generally and model-independently, the LHC allows us to test the behaviour of SM particles in a qualitatively new energy region where the electroweak symmetry is restored; studying hard-scattering events involving photons provides a way to test the validity of the Standard Model at such very high energies using clean, colorless probes that are significantly better reconstructed and calibrated than hadronic jets.

Looking back at the work done, I am proud to have played a part in one of the major discoveries in particle physics; but also, to have paved the way for complementary measurements that can shed more light on the newly observed particle or can explore the Standard Model validity in a previously unexplored energy regime, and more generally to have contributed to a better exploitation of our electromagnetic calorimeter in physics analyses not limited to the Higgs-boson search. Working on these items over the years I have had the chance and the pleasure to collaborate or discuss with a large group of bright and dedicated colleagues, either from ATLAS, CMS, or the theoretical community; I learned a lot from the more experienced ones, and tried to communicate my own knowledge and passion to the younger ones that I either had the luck to coordinate, as convener of an ATLAS analysis working group, or supervise, as director of master-level stages or co-director of an ongoing Ph.D. thesis at Université Pierre et Marie Curie (UPMC), in Paris. I like to think of my work – the techniques investigated and the results obtained – as just one piece of a larger, unfinished puzzle: time to roll up my sleeves and search for the next piece!

Bibliography

[1] BABAR Collaboration, B. Aubert et al., *Measurement of the branching fractions and CP asymmetry of $B^- \to D^0_{CP} K^-$ decays with the BABAR detector*, Phys. Rev. Lett. **92** (2004) 202002, arXiv:hep-ex/0311032 [hep-ex].

[2] BABAR Collaboration, B. Aubert et al., *Measurements of the branching fractions and CP-asymmetries of $B^- \to D^0_{(CP)} K^-$ decays*, Phys. Rev. **D73** (2006) 051105, arXiv:hep-ex/0512067 [hep-ex].

[3] BABAR Collaboration, P. del Amo Sanchez et al., *Measurement of CP observables in $B^\pm \to D_{CP} K^\pm$ decays and constraints on the CKM angle γ*, Phys. Rev. **D82** (2010) 072004, arXiv:1007.0504 [hep-ex].

[4] BABAR Collaboration, J.-P. Lees et al., *Observation of direct CP violation in the measurement of the Cabibbo-Kobayashi-Maskawa angle γ with $B^\pm \to D^{(*)} K^{(*)\pm}$ decays*, Phys. Rev. **D87** (2013) 052015, arXiv:1301.1029 [hep-ex].

[5] S. Bettarini et al., *Measurement of the charge collection efficiency after heavy non-uniform irradiation in BABAR silicon detectors*, IEEE Trans. Nucl. Sci. **52** (2005) 1054–1060.

[6] S. Bettarini et al., *The SLIM5 low mass silicon tracker demonstrator*, Nucl. Instrum. Meth. **A623** (2010) 942–953.

[7] G. Calderini, M. Benoit, N. Dinu, A. Lounis, and G. Marchiori, *Simulations of planar pixel sensors for the ATLAS high luminosity upgrade*, Nucl. Instrum. Meth. **A636** (2011) no. 1, Supplement 1, S37 – S41.

[8] J. Weingarten, S. Altenheiner, M. Beimforde, M. Benoit, M. Bomben, et al., *Planar Pixel Sensors for the ATLAS Upgrade: Beam Tests results*, JINST **7** (2012) P10028, arXiv:1204.1266 [physics.ins-det].

[9] M. Bomben, A. Bagolini, M. Boscardin, et al., *Development of edgeless n-on-p planar pixel sensors for future ATLAS upgrades*, Nucl. Instrum. Meth. **A712** (2013) no. 0, 41.

[10] ATLAS Collaboration, *Expected photon performance in the ATLAS experiment*, ATLAS-PHYS-PUB-2011-007 (2011) . http://cds.cern.ch/record/1345329.

[11] ATLAS Collaboration, *Measurements of the photon identification efficiency with the ATLAS detector using 4.9 fb^{-1} of pp collision data collected in 2011*, ATLAS-CONF-2012-123 (2012) . http://cds.cern.ch/record/1473426/.

[12] ATLAS Collaboration, *Photon identification efficiency measurements using $Z \to \ell\ell\gamma$ events in 20.7 fb^{-1} of pp collisions collected by ATLAS at 8 TeV in 2012*, 2012. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/EGAMMA/PublicPlots/20130304/ATL-COM-PHYS-2013-244.

[13] ATLAS Collaboration, *Performance of the ATLAS Electron and Photon Trigger in pp Collisions at $\sqrt{s} = 8$ TeV in 2012*, ATL-COM-DAQ-2013-121 (2013) . https://cds.cern.ch/record/1609629.

[14] ATLAS Collaboration, *Measurement of the inclusive isolated photon cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D **83** (2011) 052005, arXiv:1012.4389 [hep-ex].

[15] ATLAS Collaboration, *Measurement of the inclusive isolated prompt photon cross-section in pp collisions at $\sqrt{s} = 7$ TeV using 35 pb^{-1} of ATLAS data*, Phys. Lett. **B706** (2011) 150, arXiv:1108.0253 [hep-ex].

[16] ATLAS Collaboration, *Measurement of the production cross section of an isolated photon associated with jets in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D **85** (2012) 092014, arXiv:1203.3161 [hep-ex].

[17] ATLAS Collaboration, *Measurement of the isolated di-photon cross-section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D **85** (2012) 012003, arXiv:1107.0581 [hep-ex].

[18] ATLAS Collaboration, *Measurement of isolated-photon pair production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **1301** (2013) 086, arXiv:1211.1913 [hep-ex].

[19] ATLAS Collaboration, *ATLAS Sensitivity Prospects for Higgs Boson Production at the LHC Running at 7 TeV*, ATL-PHYS-PUB-2010-009 (2010) . https://cds.cern.ch/record/1278455.

[20] ATLAS Collaboration, *Measurement of the backgrounds to the $H \to \gamma\gamma$ search and reappraisal of its sensitivity with 37 pb^{-1} of data recorded by the ATLAS detector*, ATLAS-CONF-2011-004 (2011) . https://cds.cern.ch/record/1327196.

[21] ATLAS Collaboration, *Search for the Higgs boson in the diphoton final state with 38 pb^{-1} of data recorded by the ATLAS detector in proton-proton collisions at sqrts=7 TeV*, ATLAS-CONF-2011-025 (2011) . https://cds.cern.ch/record/1336758.

[22] ATLAS Collaboration, *Update of the Background Studies in the Search for the Higgs Boson in the Two Photons Channel in pp Collisions at $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2011-071 (2011) . http://cds.cern.ch/record/1349312.

[23] ATLAS Collaboration, *Search for the Higgs Boson in the Diphoton Channel with the ATLAS Detector using 209 pb^{-1} of 7 TeV Data taken in 2011*, ATLAS-CONF-2011-085 (2011) . http://cds.cern.ch/record/1356193.

[24] ATLAS Collaboration, *Search for the Standard Model Higgs boson in the two photon decay channel with the ATLAS detector at the LHC*, Phys. Lett. **B 705** (2011) 452, arXiv:1108.5895 [hep-ex].

[25] ATLAS Collaboration, *Search for the Standard Model Higgs boson in the diphoton decay channel with 4.9 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV with ATLAS*, Phys. Rev. Lett. **108** (2012) 111803, arXiv:1202.1414 [hep-ex].

[26] ATLAS Collaboration, *Observation of an excess of events in the search for the Standard Model Higgs boson in the $\gamma\gamma$ channel with the ATLAS detector*, ATLAS-CONF-2012-091 (2012) . https://cds.cern.ch/record/1460410.

[27] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv:1207.7214 [hep-ex].

[28] ATLAS Collaboration, *Search for the Standard Model Higgs boson in the $H \to Z\gamma$ decay mode with pp collisions at $\sqrt{s} =$ 7 and 8 TeV*, ATLAS-CONF-2013-009 (2013) . https://cds.cern.ch/record/1523683.

[29] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[30] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[31] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

[32] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[33] A. Salam, *in Elementary Particle Theory, p. 367*. Almqvist and Wiksell, Stockholm, 1968.

[34] S. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (1970) 1285.

[35] M. Gell-Mann and Y. Neeman, The Eight Fold Way,. Benjamin, New York, 1964.

[36] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.

[37] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[38] G. Guralnik, C. Hagen, and T. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.

[39] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[40] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev. **145** (1966) 1156.

[41] T. Kibble, *Symmetry breaking in non-Abelian gauge theories*, Phys. Rev. **155** (1967) 1554.

[42] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), *Handbook of LHC Higgs cross sections: 1. Inclusive observables*, CERN-2011-002 (2011) , arXiv:1101.0593 [hep-ph].

[43] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), *Handbook of LHC Higgs cross sections: 2. Differential distributions*, CERN-2012-002 (2012) , arXiv:1201.3084 [hep-ph].

[44] C. J. Seez, T. S. Virdee, L. Di Lella, R. H. Kleiss, Z. Kunszt, and W. J. Stirling, *Photon decay modes of the intermediate mass Higgs*, Proc. Large Hadron Collider Workshop, v.2 (1990) 474. https://cds.cern.ch/record/220524.

[45] L. Fayard and G. Unal, *Search for Higgs decay into photons with EAGLE*, ATL-PHYS-92-001 (1992) . https://cds.cern.ch/record/682120.

[46] C.-W. Chiang and K. Yagyu, *Higgs boson decays to $\gamma\gamma$ and $Z\gamma$ in models with Higgs extensions*, Phys. Rev. D **87** (2013) 033003, arXiv:1207.1065 [hep-ph].

[47] I. Low, J. Lykken, and G. Shaughnessy, *Singlet scalars as Higgs imposters at the Large Hadron Collider*, Phys. Rev. D **84** (2011) 035027, arXiv:1105.4587 [hep-ph].

[48] B. A. Arbuzov, *The LHC state at 125.5 GeV and FNAL data as an evidence for the existence of the new class of particles – W-hadrons*, arXiv:1209.2831 [hep-ph].

[49] B. Coleppa, K. Kumar, and H. E. Logan, *Can the 126 GeV boson be a pseudoscalar?*, Phys. Rev. D **86** (2012) 075022, arXiv:1208.2692 [hep-ph].

[50] A. Azatov, R. Contino, A. Di Iura, and J. Galloway, *New Prospects for Higgs Compositeness in $h \to Z\gamma$*, arXiv:1308.2676 [hep-ph].

[51] L. Randall and R. Sundrum, *A large mass hierarchy from a small extra dimension*, Phys. Rev. Lett. **83** (1999) 3370–3373.

[52] M. Dine, W. Fischler, and M. Srednicki, *Supersymmetric Technicolor*, Nucl. Phys. B **189** (1981) 575.

[53] S. Dimopoulos and S. Raby, *Supercolor*, Nucl. Phys. B **192** (1981) 353.

[54] C. R. Nappi and B. A. Ovrut, *Supersymmetric Extension of the $SU(3) \times SU(2) \times U(1)$ Model*, Phys. Lett. B **113** (1982) 175.

[55] S. Bhattacharya, S. S. Chauhan, B. C. Choudhary, and D. Choudhury, *Quark Excitations Through the Prism of Direct Photon Plus Jet at the LHC*, Phys. Rev. D **80** (2009) 015014, arXiv:0901.3927 [hep-ph].

[56] U. Baur, M. Spira, and P. Zerwas, *Excited quark and lepton production at hadron colliders*, Phys. Rev. D **42** (1990) 815.

[57] O. Çak ır and R. Mehdiyev, *Excited quark production at the CERN LHC*, Phys. Rev. D **60** (1999) 034004.

[58] P. Meade and L. Randall, *Black Holes and Quantum Gravity at the LHC*, JHEP **05** (2008) 003, arXiv:0708.3017 [hep-ph].

[59] X. Calmet, W. Gong, and S. D. Hsu, *Colorful quantum black holes at the LHC*, Phys. Lett. **B668** (2008) 20, arXiv:0806.4605 [hep-ph].

[60] D. M. Gingrich, *Quantum black holes with charge, colour, and spin at the LHC*, J. Phys. **G37** (2010) 105008, arXiv:0912.0826 [hep-ph].

[61] J. Kang and M. A. Luty, *Macroscopic Strings and 'Quirks' at Colliders*, JHEP **11** (2009) 065, arXiv:0805.4642 [hep-ph].

[62] S. P. Martin, *Quirks in supersymmetry with gauge coupling unification*, Phys. Rev. D **83** (2011) 035019, arXiv:1012.2072 [hep-ph].

[63] L. A. Anchordoqui, H. Goldberg, S. Nawata, and T. R. Taylor, *Direct photons as probes of low mass strings at the CERN LHC*, Phys. Rev. D **78** (2008) 016005, arXiv:0804.2013 [hep-ph].

[64] L. A. Anchordoqui, H. Goldberg, D. Lust, S. Nawata, S. Stieberger, et al., *Dijet signals for low mass strings at the LHC*, Phys. Rev. Lett. **101** (2008) 241803, arXiv:0808.0497 [hep-ph].

[65] P. Aurenche, R. Baier, M. Fontannaz, and D. Schiff, *Prompt photon production at large p_T. Scheme invariant QCD predictions and comparison with experiment*, Nucl. Phys. B **297** (1988) 661.

[66] P. Aurenche, R. Baier, M. Fontannaz, J. F. Owens, and M. Werlen, *Gluon content of the nucleon probed with real and virtual photons*, Phys. Rev. D **39** (1989) 3275.

[67] W. Vogelsang and A. Vogt, *Constraints on the proton's gluon distribution from prompt photon production*, Nucl.Phys. **B453** (1995) 334, arXiv:hep-ph/9505404 [hep-ph].

[68] L. Evans, *The Large Hadron Collider from Conception to Commissioning: A Personal Recollection*, Rev. Accl. Sci. Tech. **03** (2010) 261.

[69] *Proceedings of the ECFA-CERN Workshop on Large Hadron Collider in the LEP Tunnel*, CERN-84/10, ECFA-84.085 (1984) . http://cds.cern.ch/record/154938.

[70] ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C73** (2013) 2518, arXiv:1302.4393 [hep-ex].

[71] ATLAS Collaboration, *Luminosity Public Results*, 2013. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[72] ATLAS Collaboration, *Beam Spot Public Results*, 2013. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/BeamSpotPublicResults.

[73] ATLAS Collaboration, *ATLAS high-level trigger, data acquisition and controls: Technical design report*, CERN-LHCC-2003-022, ATLAS-TRD-016 (2003) . https://cds.cern.ch/record/616089.

[74] ATLAS Collaboration, *Data Quality Information*, 2013. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010.

[75] ATLAS Collaboration, *Expected performance of the ATLAS experiment: detector, trigger and physics, "Tracking"*, CERN-OPEN-2008-020 (2009) 16–41, arXiv:0901.0512 [hep-ex].

[76] ATLAS Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, ATLAS-CONF-2012-042 (2012) . https://cds.cern.ch/record/1435196.

[77] ATLAS Collaboration, *Estimating Track Momentum Resolution in Minimum Bias Events using Simulation and K_S in $\sqrt{s} = 900$ GeV collision data*, ATLAS-CONF-2010-009 (2010) . https://cds.cern.ch/record/1274005.

[78] ATLAS Collaboration, *Kinematic Distributions of K^0_S and Λ decays in collision data at $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2010-033 (2010) . https://cds.cern.ch/record/1277668.

[79] ATLAS Collaboration, *J/ψ Performance of the ATLAS Inner Detector*, ATLAS-CONF-2010-078 (2010) . https://cds.cern.ch/record/1281369.

[80] ATLAS Collaboration, *Tracking Results and Comparison to Monte Carlo simulation at $\sqrt{s} = 900$ GeV*, ATLAS-CONF-2010-011 (2010) . https://cds.cern.ch/record/1276323.

[81] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047 (2012), http://cds.cern.ch/record/1449796.

[82] ATLAS Collaboration, *Particle Identification Performance of the ATLAS Transition Radiation Tracker*, ATLAS-CONF-2011-128 (2011) . https://cds.cern.ch/record/1383793.

[83] ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur. Phys. J. **C72** (2012) 1909, arXiv:1110.3174 [hep-ex].

[84] ATLAS Collaboration, *Expected performance of the ATLAS experiment: detector, trigger and physics, "Muons"*, CERN-OPEN-2008-020 (2009) 161–228, arXiv:0901.0512 [hep-ex].

[85] ATLAS Collaboration, *Determination of the muon reconstruction efficiency in ATLAS at the Z resonance in proton-proton collisons at* $\sqrt{s} = 7$ *TeV*, ATLAS-CONF-2011-008 (2011) . https://cds.cern.ch/record/1330715.

[86] ATLAS Collaboration, A. Salvucci, *Measurement of muon momentum resolution of the ATLAS detector*, EPJ Web Conf. **28** (2012) 12039, arXiv:1201.4704 [physics.ins-det].

[87] ATLAS Collaboration, *MCP Momentum Scale and Z lineshape plots - 2012 Data (Moriond 2013 dataset)*, 2013. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/MUON/PublicPlots/2013/ATL-COM-PHYS-2013-340.

[88] ATLAS Collaboration, *MCP Mass Performance plots - HCP dataset 2012*, 2012. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/MUON/PublicPlots/2012/ATL-COM-MUON-2013-006.

[89] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-*k_{T} *jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[90] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at* $\sqrt{s} = 7$ *TeV*, Eur. Phys. J. **C73** (2013) 2304, arXiv:1112.6426 [hep-ex].

[91] ATLAS Collaboration, *Expected performance of the ATLAS experiment: detector, trigger and physics, "Jets and missing transverse energy"*, CERN-OPEN-2008-020 (2009) 261–396, arXiv:0901.0512 [hep-ex].

[92] A. Collaboration, *Jet energy resolution in proton-proton collisions at* $\sqrt{s} = 7$ *TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. **C73** (2013) 2306, arXiv:1210.6210 [hep-ex].

[93] ATLAS Collaboration, *Jet/Etmiss Approved 2013 JES Uncertainty*, 2013. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2013JESUncertainty.

[94] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS with 2011 Proton-Proton Collisions at sqrts = 7 TeV*, ATLAS-CONF-2012-101 (2012) . https://cds.cern.ch/record/1463915/.

[95] M. Cacciari, G. P. Salam, and G. Soyez, *The catchment area of jets*, JHEP **04** (2008) 005, arXiv:0802.1188 [hep-ph].

[96] M. Cacciari, G. P. Salam, and S. Sapeta, *On the characterisation of the underlying event*, JHEP **04** (2010) 065, arXiv:0912.4926 [hep-ph].

[97] ATLAS Collaboration, *Jet/Etmiss Approved 2013 EtMiss*, 2013. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2013EtMiss.

[98] ATLAS Collaboration, *Expected performance of the ATLAS experiment: detector, triggers and physics, "Electrons and Photons"*, arXiv:0901.0512 [hep-ex].

[99] ATLAS Collaboration, *Expected electron performance in the ATLAS experiment*, ATLAS-PHYS-PUB-2011-006 (2011) . http://cds.cern.ch/record/1345327.

[100] ATLAS Collaboration, *Electron energy response stability with pile-up in 2012 data with 13* fb^{-1}, 2012. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/EGAMMA/PublicPlots/20121129/ATL-COM-PHYS-2012-1668.

[101] ATLAS Electromagnetic Liquid Argon Calorimeter Group, B. Aubert, et al., *Performance of the ATLAS electromagnetic calorimeter end-cap module 0*, Nucl. Instrum. Meth. **A500** (2003) 178–201.

[102] ATLAS Electromagnetic Liquid Argon Calorimeter Group, B. Aubert, et al., *Performance of the ATLAS electromagnetic calorimeter barrel module 0*, Nucl. Instrum. Meth. **A500** (2003) 202–231.

[103] M. Aharrouche et al., *Response uniformity of the ATLAS liquid argon electromagnetic calorimeter*, Nucl. Instrum. Meth. **A582** (2007) no. 2, 429 – 455.

[104] ATLAS Collaboration, *Calorimeter shower shapes of photons from $Z \to \ell\ell\gamma$ decays in 20.3 fb^{-1} of pp collisions collected by ATLAS at $\sqrt{s} = 8$ TeV in 2012*, 2013. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/EGAMMA/PublicPlots/20130522/ATL-COM-PHYS-2013-600/index.html.

[105] M. Hance, D. Olivito, and H. Williams, *Performance Studies for e/gamma Calorimeter Isolation*, ATL-COM-PHYS-2011-1186 (2011) . https://cds.cern.ch/record/1379530.

[106] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. D **48** (1993) 3160–3166.

[107] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, Nucl. Phys. B **406** (1993) 187–224.

[108] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, Phys. Lett. B **641** (2006) 57–61.

[109] S. Laplace and J. de Vivie, *Calorimeter isolation and pile-up*, ATL-COM-PHYS-2012-467 (2012) . https://cds.cern.ch/record/1444890.

[110] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[111] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, *Parton Fragmentation and String Dynamics*, Phys. Rept. **97** (1983) 31–145.

[112] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4 - a simulation toolkit*, Nucl. Instrum. Methods A **506** (2003) 250.

[113] ATLAS Collaboration, *The ATLAS simulation infrastructure*, Eur. Phys. J. C **70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[114] A. Hoecker et al., *The TMVA Users Guide*, CERN-OPEN-2007-007, arXiv:physics/0703039 (2007-2009) .

[115] P. Speckmayer, A. Hocker, J. Stelzer, and H. Voss, *The toolkit for multivariate data analysis, TMVA 4*, J. Phys. Conf. Ser. **219** (2010) 032057.

[116] I. Antcheva et al., *ROOT, a C++ framework for petabyte data storage, statistical analysis and visualization*, Computer Physics Communications **180** (2009) no. 12, 2499 – 2512.

[117] K. Liu, Y. Liu, G. Marchiori, and E. Soldatov, *Measurement of the identification efficiency of isolated prompt photons using radiative $Z \to \ell\ell\gamma$ decays in 4.9 fb^{-1} of ATLAS data*, ATL-COM-PHYS-2012-382 (2012) . https://cdsweb.cern.ch/record/1437004.

[118] M. Jimenez and K. Tackmann, *Photon identification efficiency extrapolated from electrons in $Z \to e^+e^-$ decays*, ATL-COM-PHYS-2012-241 (2012) . https://cds.cern.ch/record/1428906.

[119] K. Liu, Y. Liu, and G. Marchiori, *Measurement of the identification efficiency of isolated prompt photons using the matrix method and 4.9 fb^{-1} of ATLAS data*, ATL-COM-PHYS-2012-242 (2012) . https://cdsweb.cern.ch/record/1428909.

[120] K. Liu, Y. Liu, and G. Marchiori, *Data-driven measurements of the identification efficiency of isolated prompt photons*, ATL-COM-PHYS-2011-1739 (2011) . https://cdsweb.cern.ch/record/1409480.

[121] ATLAS Collaboration, *Expected performance of the ATLAS experiment: detector, trigger and physics, "Cross-Sections, Monte Carlo Simulations and Systematic Uncertainties"*, CERN-OPEN-2008-020 (2009) 13–14, arXiv:0901.0512 [hep-ex].

[122] ATLAS Collaboration, *Measurements of the properties of the Higgs-like boson in the two photon decay channel with the ATLAS detector using 25* fb^{-1} *of proton-proton collision data*, ATLAS-CONF-2013-012 (2013) . https://cds.cern.ch/record/1523698.

[123] L. Yuan, G. Calderini, S. De Cecco, S. Jin, B. Laforge, G. Marchiori, J. Ocariz, L. Roos, H. Torres, and N. Trinh, *Data-driven methods for photon trigger efficiency measurement*, ATL-COM-PHYS-2010-113 (2010) . https://cds.cern.ch/record/1245866/.

[124] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv:0811.4622 [hep-ph].

[125] S. Hoeche, S. Schumann, and F. Siegert, *Hard photon production and matrix-element parton-shower merging*, Phys. Rev. D **81** (2010) 034026, arXiv:0912.3501 [hep-ph].

[126] ATLAS Collaboration, *Evidence for prompt photon production in pp collisions at* $\sqrt{s} = 7$ TeV *with the ATLAS detector*, ATLAS-CONF-2010-077 (2010) . http://cds.cern.ch/record/1281368.

[127] ATLAS Collaboration, *Inclusive cross sections of isolated prompt photons in pp collisions at* $\sqrt{s} = 7$ *TeV measured with the ATLAS detector using 4.7* fb^{-1}, ATLAS-CONF-2013-022 (2013) . https://cds.cern.ch/record/1525723/.

[128] G. R. Farrar, *Experimental Means for Distinguishing Models of Large* p_T *Inclusive Scattering*, Phys. Lett. **B67** (1977) 337.

[129] P. Aurenche, A. Douiri, R. Baier, M. Fontannaz, and D. Schiff, *Prompt Photon Production at Large* p_T *in QCD Beyond the Leading Order*, Phys. Lett. **B140** (1984) 87.

[130] J. Owens, *Large Momentum Transfer Production of Direct Photons, Jets, and Particles*, Rev.Mod.Phys. **59** (1987) 465.

[131] S. Catani et al., *Cross section of isolated prompt photons in hadron-hadron collisions*, JHEP **05** (2002) 028, arXiv:hep-ph/0204023.

[132] M. Klein and R. Yoshida, *Collider Physics at HERA*, Prog. Part. Nucl. Phys. **61** (2008) 343, arXiv:0805.3334 [hep-ex].

[133] R. Ichou and D. d'Enterria, *Sensitivity of isolated photon production at TeV hadron colliders to the gluon distribution in the proton*, Phys. Rev. **D82** (2010) 014015, arXiv:1005.4529 [hep-ph].

[134] T. Binoth, J. P. Guillet, E. Pilon, and M. Werlen, *A full next-to-leading order study of direct photon pair production in hadronic collisions*, Eur. Phys. J. C **16** (2000) 311, arXiv:hep-ph/9911340 [hep-ph].

[135] P. Aurenche et al., *Recent critical study of photon production in hadronic collisions*, Phys. Rev. D **73** (2006) 094007, arXiv:hep-ph/0602133.

[136] Z. Belghobsi, M. Fontannaz, J.-P. Guillet, G. Heinrich, E. Pilon, et al., *Photon - Jet Correlations and Constraints on Fragmentation Functions*, Phys. Rev. D **79** (2009) 114024, arXiv:0903.4834 [hep-ph].

[137] L. Ametller, E. Gava, N. Paver, and D. Treleani, *Role of the QCD induced gluon-gluon coupling to gauge boson pairs in the multi-TeV region*, Phys. Rev. D **32** (1985) 1699.

[138] L. Bourhis, M. Fontannaz, and J. Guillet, *Quarks and gluon fragmentation functions into photons*, Eur.Phys.J. **C2** (1998) 529–537, arXiv:hep-ph/9704447 [hep-ph].

[139] R. M. Whalley, D. Bourilkov, and R. Group, *The Les Houches accord PDFs (LHAPDF) and LHAGLUE*, arXiv:hep-ph/0508110.

[140] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012.

[141] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[142] Z. Bern, L. Dixon, and C. Schmidt, *Isolating a light Higgs boson from the diphoton background at the LHC*, Phys. Rev. D **66** (2002) 074018, arXiv:hep-ph/0206194 [hep-ph].

[143] S. Catani, L. Cieri, D. de Florian, G. Ferrera, and M. Grazzini, *Diphoton production at hadron colliders: a fully-differential QCD calculation at NNLO*, Phys. Rev. Lett. **108** (2012) 072001, arXiv:1110.2375 [hep-ph].

[144] S. Frixione, *Isolated photons in perturbative QCD*, Phys.Lett. **B429** (1998) 369–374, arXiv:hep-ph/9801442 [hep-ph].

[145] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C 63** (2009) 189, arXiv:0901.0002 [hep-ph].

[146] A. Sherstnev and R. S. Thorne, *Parton distributions for LO generators*, Eur. Phys. J. C **55** (2008) 553, arXiv:0711.2473 [hep-ph].

[147] G. Corcella et al., *HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including super-symmetric processes)* , JHEP **01** (2001) 010.

[148] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. **C 72** (1996) 637, arXiv:hep-ph/9601371.

[149] G. D'Agostini, *A multidimensional unfolding method based on Bayes' theorem*, Nucl. Instrum. Methods A **362** (1995) 487.

[150] A. Hoecker and V. Kartvelishvili, *SVD approach to data unfolding*, Nucl. Instrum. Methods A **372** (1996) no. 3, 469.

[151] D0 Collaboration, V. M. Abazov et al., *Measurement of the isolated photon cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Lett. B **639** (2006) 151, arXiv:hep-ex/0511054 [hep-ex].

[152] CDF Collaboration, T. Aaltonen et al., *Measurement of the inclusive isolated prompt photon cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using the CDF detector*, Phys. Rev. D **80** (2009) 111106(R), arXiv:0910.3623 [hep-ex].

[153] H1 Collaboration, D. Aaron et al., *Prompt photons in photoproduction at HERA*, Eur. Phys. J. C **66** (2010) 17, arXiv:0910.5631 [hep-ex].

[154] ZEUS Collaboration, J. Breitweg et al., *Measurement of inclusive prompt photon photoproduction at HERA*, Phys. Lett. B **472** (2000) 175–188, arXiv:hep-ex/9910045.

[155] CMS Collaboration, *Measurement of the Differential Cross Section for Isolated Prompt Photon Production in pp Collisions at 7 TeV*, Phys. Rev. D **84** (2011) 052011, arXiv:1108.2044 [hep-ex].

[156] A. D. Martin, R. Roberts, W. J. Stirling, and R. Thorne, *Parton distributions and the LHC: W and Z production*, Eur. Phys. J. **C14** (2000) 133, arXiv:hep-ph/9907231 [hep-ph].

[157] E706 Collaboration, L. Apanasevich et al., *Evidence for parton k_T effects in high p_T particle production*, Phys. Rev. Lett. **81** (1998) 2642–2645, arXiv:hep-ex/9711017 [hep-ex].

[158] D. d'Enterria and J. Rojo, *Quantitative constraints on the gluon distribution function in the proton from collider isolated-photon data*, Nucl. Phys. **B860** (2012) 311, arXiv:1202.1762 [hep-ph].

[159] NNPDF Collaboration Collaboration, R. D. Ball et al., *Reweighting NNPDFs: the W lepton asymmetry*, Nucl.Phys. **B849** (2011) 112–143, arXiv:1012.0836 [hep-ph].

[160] D0 Collaboration, V. M. Abazov et al., *Measurement of the differential cross-section for the production of an isolated photon with associated jet in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Lett. B **666** (2008) 435, arXiv:0804.1107 [hep-ex].

[161] CDF Collaboration, F. Abe et al., *Properties of photon plus two jet events in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. D **57** (1998) 67.

[162] Axial Field Spectrometer Collaboration, T. Akesson et al., *Direct photon plus away-side jet production in pp collisions at $\sqrt{s} = 63$ GeV and a determination of the gluon distribution*, Z. Phys. C **34** (1987) 293.

[163] UA2 Collaboration, J. Alitti et al., *Measurement of the gluon structure function from direct photon data at the CERN anti-p p collider*, Phys. Lett. B **299** (1993) 174.

[164] H1 Collaboration, F. D. Aaron et al., *Measurement of isolated photon production in deep-inelastic scattering at HERA*, Eur. Phys. J. C **54** (2008) 371, arXiv:0711.4578 [hep-ex].

[165] ZEUS Collaboration, S. Chekanov et al., *Measurement of prompt photons with associated jets in photoproduction at HERA*, Eur. Phys. J. **C49** (2007) 511, arXiv:hep-ex/0608028.

[166] ZEUS Collaboration, S. Chekanov et al., *Observation of isolated high-E_T photons in deep inelastic scattering*, Phys. Lett. B **595** (2004) 86, arXiv:hep-ex/0402019.

[167] CMS Collaboration, *Measurement of triple-differential cross section of γ+jet production*, CMS-PAS-QCD-11-005 (2013) . https://cds.cern.ch/record/1525534/.

[168] D0 Collaboration, V. M. Abazov et al., *Measurement of the differential cross section of photon plus jet production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, arXiv:1308.2708 [hep-ex].

[169] L. Carminati, G. Costa, D. D'Enterria, I. Koletsou, G. Marchiori, et al., *Sensitivity of the LHC isolated-γ+jet data to the parton distribution functions of the proton*, EPL **101** (2013) 61002, arXiv:1212.5511 [hep-ph].

[170] J. C. Collins and D. E. Soper, *Angular distribution of dileptons in high-energy hadron collisions*, Phys. Rev. D **16** (1977) 2219.

[171] S. Schmitt, *TUnfold: an algorithm for correcting migration effects in high energy physics*, JINST **7** (2012) T10003, arXiv:1205.6201 [physics.data-an].

[172] T. Binoth, J. P. Guillet, E. Pilon, and M. Werlen, *Beyond leading order effects in photon pair production at the Fermilab Tevatron*, Phys. Rev. D **63** (2001) 114016, arXiv:hep-ph/0012191 [hep-ph].

[173] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv:1207.7235 [hep-ex].

[174] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: A program for Higgs boson decays in the standard model and its supersymmetric extension*, Comput. Phys. Commun. **108** (1998) 56, arXiv:hep-ph/9704448.

[175] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, *Precise predictions for the Higgs-boson decay $H \to WW/ZZ \to 4$leptons*, Phys. Rev. **D 74** (2006) 013004, arXiv:hep-ph/0604011.

[176] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, *NNLO computational techniques: the cases $H \to \gamma\gamma$ and $H \to gg$*, Nucl. Phys. **B 811** (2009) 182–273, arXiv:0809.3667 [hep-ph].

[177] A. Djouadi, M. Spira, and P. M. Zerwas, *Production of Higgs bosons in proton colliders: QCD corrections*, Phys. Lett. **B 264** (1991) 440–446.

[178] S. Dawson, *Radiative corrections to Higgs boson production*, Nucl. Phys. **B 359** (1991) 283–300.

[179] M. Spira, A. Djouadi, D. Graudenz, and P. M. Zerwas, *Higgs boson production at the LHC*, Nucl. Phys. **B 453** (1995) 17, arXiv:hep-ph/9504378.

[180] R. V. Harlander and W. B. Kilgore, *Next-to-next-to-leading order Higgs production at hadron colliders*, Phys. Rev. Lett. **88** (2002) 201801, arXiv:hep-ph/0201206.

[181] C. Anastasiou and K. Melnikov, *Higgs boson production at hadron colliders in NNLO QCD*, Nucl. Phys. **B 646** (2002) 220, arXiv:hep-ph/0207004.

[182] V. Ravindran, J. Smith, and W. L. van Neerven, *NNLO corrections to the total cross section for Higgs boson production in hadron hadron collisions*, Nucl. Phys. **B 665** (2003) 325–366, arXiv:hep-ph/0302135.

[183] U. Aglietti, R. Bonciani, G. Degrassi, and A. Vicini, *Two-loop light fermion contribution to Higgs production and decays*, Phys. Lett. **B 595** (2004) 432–441, arXiv:hep-ph/0404071.

[184] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, *NLO electroweak corrections to Higgs boson production at hadron colliders*, Phys. Lett. **B 670** (2008) 12–17, arXiv:0809.1301 [hep-ph].

[185] D. de Florian and M. Grazzini, *Higgs production at the LHC: updated cross sections at $\sqrt{s} = 8$ TeV*, Phys. Lett. **B718** (2012) 117, arXiv:1206.4133 [hep-ph].

[186] C. Anastasiou, S. Buehler, F. Herzog, and A. Lazopoulos, *Inclusive Higgs boson cross-section for the LHC at 8 TeV*, JHEP **04** (2012) 004, arXiv:1202.3638 [hep-ph].

[187] J. Baglio and A. Djouadi, *Higgs production at the lHC*, JHEP **1103** (2011) 055, arXiv:1012.0530 [hep-ph].

[188] M. Ciccolini, A. Denner, and S. Dittmaier, *Strong and electroweak corrections to the production of Higgs+2jets via weak interactions at the LHC*, Phys. Rev. Lett. **99** (2007) 161803, arXiv:0707.0381 [hep-ph].

[189] M. Ciccolini, A. Denner, and S. Dittmaier, *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC*, Phys. Rev. D **77** (2008) 013002, arXiv:0710.4749 [hep-ph].

[190] K. Arnold et al., *VBFNLO: A parton level Monte Carlo for processes with electroweak bosons*, Comput. Phys. Commun. **180** (2009) 1661–1670, arXiv:0811.4559 [hep-ph].

[191] P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro, *Higgs production via vector-boson fusion at NNLO in QCD*, Phys. Rev. Lett. **105** (2010) 011801, arXiv:1003.4451 [hep-ph].

[192] T. Han and S. Willenbrock, *QCD correction to the $pp \to WH$ and ZH total cross- sections*, Phys. Lett. **B 273** (1991) 167–172.

[193] O. Brein, A. Djouadi, and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, Phys. Lett. **B 579** (2004) 149, arXiv:hep-ph/0307206.

[194] M. L. Ciccolini, S. Dittmaier, and M. Kramer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, Phys. Rev. D **68** (2003) 073003, arXiv:hep-ph/0306234.

[195] W. Beenakker, S. Dittmaier, M. Kramer, B. Plumper, M. Spira, et al., *Higgs radiation off top quarks at the Tevatron and the LHC*, Phys. Rev. Lett. **87** (2001) 201805, arXiv:hep-ph/0107081 [hep-ph].

[196] W. Beenakker et al., *NLO QCD corrections to $t\bar{t}H$ production in hadron collisions.* , Nucl. Phys. **B 653** (2003) 151–203, arXiv:hep-ph/0211352.

[197] S. Dawson, L. Orr, L. Reina, and D. Wackeroth, *Next-to-leading order QCD corrections to $pp \to t\bar{t}h$ at the CERN Large Hadron Collider*, Phys. Rev. D **67** (2003) 071503, arXiv:hep-ph/0211438 [hep-ph].

[198] S. Dawson, C. Jackson, L. H. Orr, L. Reina, and D. Wackeroth, *Associated Higgs production with top quarks at the Large Hadron Collider: NLO QCD corrections*, Phys. Rev. D **68** (2003) 034022, arXiv:hep-ph/0305087.

[199] ATLAS and CMS Collaborations, *Procedure for the LHC Higgs boson search combination in Summer 2011*, ATLAS-PHYS-PUB-2011-011, CMS NOTE-2011/005 (2011) . https://cds.cern.ch/record/1375842.

[200] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C 71** (2011) 1554, arXiv:1007.1727 [physics.data-an].

[201] A. L. Read, *Presentation of search results: The CL_s technique*, J. Phys. G: Nucl. Part. Phys **28** (2002) 2693.

[202] ATLAS Collaboration, *ATLAS: technical proposal for a general-purpose pp experiment at the Large Hadron Collider at CERN*, CERN-LHCC-94-43 (1994) . http://cds.cern.ch/record/290968.

[203] ATLAS Collaboration, *Expected performance of the ATLAS experiment: detector, trigger and physics, "Prospects for the Discovery of the Standard Model Higgs Boson Using the $H \to \gamma\gamma$ Decay"*, CERN-OPEN-2008-020 (2009) 1212–1242, arXiv:0901.0512 [hep-ex].

[204] S. Myers and F. Zimmermann, *Summary*, Proceedings of Chamonix 2009 workshop on LHC Performance (2009) . https://cds.cern.ch/record/1172838.

[205] S. Myers and F. Zimmermann, *Summary*, Proceedings of Chamonix 2010 workshop on LHC Performance (2010) . http://cds.cern.ch/record/1236824.

[206] H. Abreu et al., *Prospects for the exclusion of a SM Higgs decaying to two photons*, ATL-PHYS-INT-2010-061 (2010) . https://cds.cern.ch/record/1270178.

[207] M. L. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001, arXiv:hep-ph/0206293.

[208] M. Oreglia, *Ph.D. thesis*, SLAC-R-0236 (1980) Appendix D.

[209] CDF Collaboration, *Search for a SM Higgs Boson with the Diphoton Final State at CDF*, CDF/PUB/EXOTIC/PUBLIC/10065 (2010) . http://www-cdf.fnal.gov/physics/new/hdg/results/hgamgam_jan10/cdf10065_HiggsGamGam54Public.pdf.

[210] D0 Collaboration, *Search for the SM Higgs Boson in gamma-gamma final states at D0 with* $L = 4.2\ fb^{-1}$ *data*, D0 Note 5858-Conf (2009) . http://www-d0.fnal.gov/Run2Physics/WWW/results/prelim/HIGGS/H66/H66.pdf.

[211] S. Alioli, P. Nason, C. Oleari and E. Re, *NLO Higgs boson production via gluon fusion matched with shower in POWHEG*, JHEP **04** (2009) 002, arXiv:0812.0578 [hep-ph].

[212] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP **02** (2010) 037, arXiv:0911.5299 [hep-ph].

[213] E. Gross and O. Vitells, *Trial factors for the look elsewhere effect in high energy physics*, Eur. Phys. J. **C 70** (2010) 525, arXiv:1005.1891 [physics.data-an].

[214] I. Low, J. Lykken, and G. Shaughnessy, *Have we observed the Higgs (imposter)?*, Phys. Rev. D **86** (2012) 093012, arXiv:1207.1093 [hep-ph].

[215] Particle Data Group, J. Beringer, et al., *Review of particle physics*, Phys. Rev. D **86** (2012) 010001.

[216] L.-B. Chen, C.-F. Qiao, and R.-L. Zhu, *Reconstructing the 125 GeV SM Higgs boson through* $\ell\bar{\ell}\gamma$, arXiv:1211.6058 [hep-ph].

[217] ATLAS Collaboration, *Luminosity determination in pp collisions at* $\sqrt{s} = 7$ TeV *using the ATLAS detector at the LHC*, Eur. Phys. J. C **71** (2011) 1630, arXiv:1101.2185 [hep-ex].

[218] A. Firan and R. Stroynowski, *Internal conversions in Higgs decays to two photons*, Phys. Rev. D **76** (2007) 057301, arXiv:0704.3987 [hep-ph].

[219] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, arXiv:hep-ph/0204244 [hep-ph].

[220] S. Frixione, P. Nason, and B. R. Webber, *Matching NLO QCD and parton showers in heavy flavour production*, JHEP **08** (2003) 007, arXiv:hep-ph/0305252.

[221] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C 45** (2006) 97–107, arXiv:hep-ph/0506026.